2018

Quantification of Understory Fuels in the Superior National Forest Using Lidar Data

Jeffrey R. Irwin
South Dakota State University

QUANTIFICATION OF UNDERSTORY FUELS IN THE SUPERIOR NATIONAL FOREST USING LIDAR DATA

BY

JEFFREY R. IRWIN

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Geography

South Dakota State University

2018

QUANTIFICATION OF UNDERSTORY FUELS IN THE SUPERIOR NATIONAL FOREST USING LIDAR DATA

JEFFREY R. IRWIN

This thesis is approved as a creditable and independent investigation by a candidate for the Master of Science degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this thesis does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Darrell Napton, Ph.D.
Thesis Advisor — Date

Robert Watrel, Ph.D.
Head, Department of Geography — Date

Kinchel Doerner, Ph.D.
Dean, Graduate School — Date

ACKNOWLEDGEMENTS

First of all, I would like to thank Darrell Napton for providing me with great advice over the last several years. Darrell was my advisor long before he was my advisor. Bruce Millett has also been a wonderful source of advice and guidance over the years. I can't express enough gratitude to Kurtis Nelson and Birgit Peterson from the USGS EROS Center. Kurtis and Birgit have given me the incredible opportunity to get my start with the USGS and have provided many great suggestions when I needed them in the process of completing this work. Birgit deserves special thanks for joining the thesis committee. I would also like to everyone from the EROS Science and Applications Branch, especially Bruce Wylie, Charlie Trautwein, Kristi Sayler, Jeff Danielson, Dean Gesch, Dean Mierau, and Michelle Knuppe. Lastly, I would like to thank my friends and family. They have been sources of inspiration and support. I particularly need to thank my wife, Rebecca, for her patience and understanding.

CONTENTS

ABBREVIATIONS……vi

LIST OF FIGURES……...…vii

LIST OF TABLES……xiv

ABSTRACT……...xv

INTRODUCTION……...1

Objectives……2

Study Area……...2

LITERATURE REVIEW……7

A Brief History of the Superior National Forest Region……..……...7

Superior National Forest Ecosystems……..9

Fire and Fuels in the Superior National Forest Region……..12

Balsam Fir Characteristics……21

Lidar……...26

UNDERSTORY COVER MODEL DEVLEOPMENT……32

Introduction……32

Materials and Methods…….....32

Results and Discussion……..48

Summary and Conclusion……..80

COMPARISON OF THE UNDERSTORY COVER MODEL TO ANCILARY DATA SETS…….81

Introduction……81

Materials and Methods……...81

CONTENTS (CONTINUED)

Results and Discussion……..89

Summary and Conclusion……..97

BALSAM FIR UNDERSTORY COVER MODEL DEVELOPMENT……...99

Introduction……99

Materials and Methods…….....99

Results and Discussion……102

Summary and Conclusion……113

ADDITIONAL DEVELOPMENT OF UNDERSTORY COVER MODELS……114

Introduction……..114

Materials and Methods…….....115

Results and Discussion……119

Summary and Conclusion……137

OVERALL SUMMARY AND CONCLUSIONS……..139

Summary……..139

Conclusions……..140

Future Work……141

REFERENCES……142

ABBREVIATIONS

CBH……..……crown/canopy base height

CFBH…….crown fuel base height

dbh……...tree diameter at breast height

EROS……...Earth Resources Observation and Science Center

ft……foot/feet

lidar…….light detection and ranging

LCBH…….live crown base height

m..…….meter(s)

MNDNR…….....Minnesota Department of Natural Resources

MTBS……...Monitoring Trends in Burn Severity

NLCD……...National Land Cover Database

r……...correlation coefficient

r^2……coefficient of determination

R^2…….coefficient of multiple determination

USFS…….United States Forest Service

USGS…….United States Geological Survey

LIST OF FIGURES

Figure 1. Map showing the location of Superior National Forest……...…3

Figure 2. Superior National Forest and the Boundary Waters Canoe Area Wilderness….4

Figure 3. The location of the study area within Superior National Forest……..5

Figure 4. Balsam fir growing in the understory of a red and white pine stand in the George Washington Pines area of Superior National Forest. Photo taken by Birgit Peterson, USGS EROS, in May 2017……22

Figure 5. Visual representation of a line intercept transect (Lutes et al. 2006)……34

Figure 6. Diagrams demonstrating the process for measuring the horizontal distribution of the canopy (Lutes et al. 2006)……...35

Figure 7. The locations of the field plots……..36

Figure 8. Phase 2 plot locations……37

Figure 9. Locations of the validation plots……...38

Figure 10. Photo taken along transect 1 of plot 27 by Kurtis Nelson, USGS EROS, July 2016. Note the dense birch and balsam fir understory……..39

Figure 11. Photo taken along transect 2 of plot 31 by Birgit Peterson, USGS EROS, May 2017. Note the dead lower limbs on the balsam fir, which add to fuel loads……...40

Figure 12. Cook County, Minnesota……41

Figure 13. A portion of the original point cloud with the points colored by elevation. Blue points are lowest in elevation and red points are highest……..42

Figure 14. A profile view from the original point cloud. Points are colored by elevation……43

Figure 15. A profile view from the normalized point cloud, notice the flat ground surface. Points are colored by elevation……...43

Figure 16. Raster image for the 99th height percentile lidar metric. Output is in meters……53

Figure 17. A graph showing the correlation of each lidar metric with the measured understory canopy. The y-axis displays the r^2 values and the x-axis displays the lidar metrics……....54

LIST OF FIGURES (CONTINUED)

Figure 18. Output from R (R Core Team 2016), showing the linear model results comparing the measured understory canopy with the relative density between 2 and 4 meters lidar metric…….55

Figure 19. Measured understory canopy vs relative density between 2 and 4 meters…..55

Figure 20. R code and output for the trial and error understory canopy model……56

Figure 21. Leaps package (Lumley 2009) graph used for predictor variable selection…57

Figure 22. Output for the five identified predictor variables……58

Figure 23. Results from removing the relative density between 4 and 6 meters metric…59

Figure 24. Output after removing the cover greater than 12 meters metric……..60

Figure 25. Leaps graph for three variable selection, confirming the final results of the five-variable effort…….60

Figure 26. Raster for the height at which 50% of the lidar points are returned metric at 30-meter resolution……61

Figure 27. Raster for the relative density between the heights of 8 and 10 meters metric at 30-meter resolution……62

Figure 28. Raster for the total cover lidar metric at 30-meter resolution…….63

Figure 29. Spatial output of the understory canopy model at 30-meter resolution……..64

Figure 30. The model expressed as understory cover at 30-meter resolution……..65

Figure 31. Statistics and histogram of the pixel values for the understory cover model..65

Figure 32. The 2006 Cavity Lake, Famine Lake, and Redeye fires and 2007 Ham Lake fire burn severity maps (Monitoring Trends in Burn Severity 2016e, Monitoring Trends in Burn Severity 2016f) displayed with the understory cover model. Green areas represent low burn severity and red areas represent high burn severity……67

Figure 33. LANDFIRE disturbance data appears to match with many of the areas that are low in understory cover…….68

Figure 34. The "cleaned up" version of the understory cover model……...69

LIST OF FIGURES (CONTINUED)

Figure 35. The pixel statistics and histogram for the understory cover model after trimming the spatial extent of the model……...70

Figure 36. The raster product after adjusting the extreme high and low values……...72

Figure 37. Statistics from the raster with the adjusted high and low values…….72

Figure 38. Rescaled understory cover model……73

Figure 39. Statistics and histogram for the rescaled understory cover model……..74

Figure 40. Graphs displaying the relationships between the measured and predicted understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory cover minus the predicted understory cover…….78

Figure 41. Map of forest types provided by the researchers from the University of Minnesota……...82

Figure 42. Map of crown fuel base height provided by the researchers from the University of Minnesota……83

Figure 43. Map of live crown base height provided by the researchers from the University of Minnesota……....84

Figure 44. The binary raster displaying the areas where the understory cover is greater than 70%.....85

Figure 45. The understory cover model clipped to the study area……86

Figure 46. National Land Cover Database land cover map for the study area. (Land cover types: 11 = Open Water; 21 = Developed, Open Space; 22 = Developed, Low Intensity; 23 = Developed, Medium Intensity; 24 = Developed, High Intensity; 31 = Barren Land (Rock/Sand/Clay); 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 81 = Pasture/Hay; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands)……88

Figure 47. Comparison of the understory cover model to the forest type map. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands, 111 = roads, 112 = water)……89

LIST OF FIGURES (CONTINUED)

Figure 48. Comparison of the understory cover model to the forest type map, results have been normalized. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands, 111 = roads)…….90

Figure 49. (Left) and Figure 50. (Right). Scatter plots showing the relationship between the understory cover model and the crown fuel base height (Figure 47) and live crown base height (Figure 48) models……..91

Figure 51. (Left). and Figure 52. (Right). Comparison of the crown fuel base height model to the forest type model. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands. Heights are in meters)……..92

Figure 53. (Left) and Figure 54. (Right). Comparison of the live crown base height model to the forest type model. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands. Heights are in meters)……..93

Figure 55. (Left). and Figure 56. (Right). Comparison of the understory cover model to 2011 land cover data from the National Land Cover Database. Class areas have been normalized in Figure 54. (Land cover classes: 11 = Open Water; 21 = Developed, Open Space; 22 = Developed, Low Intensity; 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands)……..94

Figure 57. (Left). and Figure 58. (Right). Comparison of the crown fuel base height model to land cover data from the National Land Cover Database. (Land cover classes: 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands)……...95

Figure 59. (Left). and Figure 60. (Right). Comparison of the live crown base height model to land cover data from the National Land Cover Database. (Land cover classes: 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands)……...96

Figure 61. The balsam fir basal area, measured in square meters per pixel (Wolter 2016b)……..100

Figure 62. Total basal area spatial model (Wolter 2016b). Basal area is in square meters per pixel……...101

LIST OF FIGURES (CONTINUED)

Figure 63. The class 5 balsam fir binary raster……...103

Figure 64. (Left) and Figure 65. (Right). Comparison of the raw (Figure 62) and normalized (Figure 63) amount of area in each of the five-balsam fir basal area classes……..103

Figure 66. The final proportion of balsam fir raster after having been clipped to the extent of Cook County and having the pixels with values greater than 1 removed…….105

Figure 67. Statistics for the final version of the proportion of balsam fir raster…….105

Figure 68. The raster product of the understory cover model and the proportion of balsam fir raster……106

Figure 69. Statistics generated from the balsam fir understory cover raster…..……107

Figure 70. The raster product of the proportion of balsam fir raster and the rescaled understory cover model……108

Figure 71. Statistics from the raster product of the rescaled understory cover raster and the proportion of balsam fir raster……...108

Figure 72. Graphs displaying the relationships between the measured and predicted balsam fir understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured balsam fir understory cover minus the predicted balsam fir understory cover……...110

Figure 73. Graphs displaying the relationships between the measured and predicted rescaled balsam fir understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured balsam fir understory cover minus the predicted rescaled balsam fir understory cover…….111

Figure 74 (Left) and Figure 75 (Right). Figure 74 displays a histogram of the measured plot understory cover for all 70 plots. Figure 75 displays a boxplot of the same data...120

Figure 76. A number line showing the mean plot understory for each of the 70 plots...120

LIST OF FIGURES (CONTINUED)

Figure 77. Graphs displaying the relationships between the measured and predicted understory cover values for the stratified random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……...122

Figure 78. Graphs displaying the relationships between the measured and predicted understory cover values for the stratified 75% of medium and 35% of high and low plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……..123

Figure 79. Graphs displaying the relationships between the measured and predicted understory cover values for the 80% training and 20% test plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……125

Figure 80. Graphs displaying the relationships between the measured and predicted understory cover values for the 80% training and 20% test plots, withholding plot 62 from the training data, random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……..126

LIST OF FIGURES (CONTINUED)

Figure 81. Graphs displaying the relationships between the measured and predicted understory cover values for the 70% training and 30% test plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……128

Figure 82. Graphs displaying the relationships between the measured and predicted understory cover values for the validation data for the model built from the 4 transect plots. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……...…129

Figure 83. Graphs displaying the relationships between the measured and predicted understory cover values at each validation plot for the revised original model containing plot 121 in the training data. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover……131

Figure 84. The understory cover model had good predictive power in several validation plots, including plots 23, 110, and 167. Plot 23 had a measured understory cover of 107.37% and a predicted understory cover of 107.87%. Plot 110 had a measured understory cover of 155.51% and a predicted understory cover of 155.13%. Plot 167 had a measured understory cover of 81.12% and a predicted understory cover of 93.44%...141

LIST OF TABLES

Table 1. The cover line intercept tree classification scheme, adapted from Lutes et al. (2006) (dbh = diameter at breast height)……...…36

Table 2. A list of calculations made for each tree species in each transect……..38

Table 3. Tree size classes (dbh = diameter at breast height)……39

Table 4. Lidar metrics derived from the normalized point cloud. The shaded metrics were used as predictor variables in regression analysis…….44

Table 5. Plot totals of understory cover measurements. (U. Canopy = the understory canopy measured in meters, plots highlighted in green were measured by the EROS team, plots highlighted in yellow were measured by the U.S. Forest Service, plots highlighted in red were excluded from the analysis, "-" indicates the transect was not measured.)…….49-52

Table 6. Plot totals of understory cover measurements for the validation plots. (U. Canopy = the understory canopy measured in meters, - indicates the transect was not measured.)……...……..75-76

Table 7. The measured and predicted understory values for the 24 validation plots. (Meas. U.C. = Measured Understory Cover, Pred. U.C. = Predicted Understory Cover)…….77

Table 8. List of the binary rasters generated for the comparison of the understory cover model with the ancillary data sets……..84

Table 9. The binary rasters produced from the crown fuel base height (CFBH) and the live crown base height (LCBH) data sets. Heights are in meters…….87

Table 10. Natural breaks (Jenks) classes generated from the balsam fir basal area raster…….102

Table 11. Plot validation data for the balsam fir understory cover models……109

Table 12. Plots in each division of understory cover……..121

Table 13. Pertinent information about each of the eight models built in this investigation. One transect plots are colored in red……132-135

ABSTRACT

QUANTIFICATION OF UNDERSTORY FUELS IN THE SUPERIOR NATIONAL FOREST USING LIDAR DATA

JEFFREY R. IRWIN

2018

Fire is a natural part of the ecosystems in the Great Lakes region. Several factors including harvesting, insect outbreaks, and fire suppression have had an impact on these ecosystems. Of particular concern is the rise in the proportion of shade tolerant species, such as balsam fir. Because of its resinous bark and easily ignited needles, balsam fir is flammable. Additionally, balsam fir is vulnerable to spruce budworm infestation, which leads to additional fuel loading from needle cast. Detecting increases in fire risk is important to fire managers in Superior National Forest for many reasons. Determining the amount and extent of understory fuels, however, is difficult, because of the area's remoteness and to the cost of collecting field data. To quantify and map the extent of understory fuels, lidar data were used in combination with field data collected using the cover line intercept method at 46 plots. Multiple lidar metrics were calculated from the original point cloud, to be used as predictor variables. Preliminary results indicated that there was a relationship ($R^2 \approx 0.65$) between the amount of understory canopy measured and the following lidar metrics: the height at which 50 percent of lidar pulses are returned, relative point density between the heights of 8 and 10 meters, and the total cover. The results were used to produce a preliminary understory cover model. Comparison of the model results with ancillary data increased confidence in the model. Cover line intercept data were collected at an additional 24 validation plots. The

validation data were compared to the model's predicted understory cover. The results produced an $R^2 \approx 0.47$, a mean difference ≈ 14.3%, and a mean absolute difference ≈ 47.0%. The understory cover model was combined with data mapping the amount of balsam fir basal area to produce a preliminary balsam fir understory cover map. This research indicates that it may be possible to map understory fuels with airborne lidar and that with proper data, individual understory components may be extracted.

INTRODUCTION

Fire is a natural part of the ecosystems within the Great Lakes region and hence indigenous species are adapted to fire. Several elements, including fire exclusion, logging practices, and insect infestations have made impacts upon those ecosystems (Blais 1983b, Buell and Niering 1957, Heinselman 1973, Heinselman 1996). One result of these impacts is the increase of balsam fir, a shade tolerant species that typically grows in the understory, on the landscape (Corace III et al. 2012, Heinselman 1973). Balsam fir is highly flammable as a result of its resinous bark and easily ignited needles (Uchytil 1991). Additionally, balsam is a ladder fuel that allows fire to spread from the forest floor into the upper canopy, initiating crown fires (Abbas et al. 2011, Heinselman 1973). Furthermore, balsam fir is prone to outbreaks of spruce budworm, which often kills trees after 3 – 5 years of perpetual defoliation. This defoliation increases fuel loads by providing cast needles and dead aerial fuel (Uchytil 1991).

The increase in balsam fir on the landscape is an important issue for forest managers in northeastern Minnesota. Even though Sturtevant et al. (2012) determined that spruce budworm outbreaks do not raise the long-term risks of wildland fire, Stocks (1987) found that fire potential was at its highest 5- 8 years after balsam fir death from spruce budworm and that fires could behave explosively. Fleming, Candau, and McAlpine (2002) determined that wildfire most frequently occurred 3 – 9 years after a spruce budworm outbreak.

Fire managers in northeastern Minnesota need to protect structures and the wildland-urban interface and this increased fire risk is of concern. Unfortunately, the amount and spatial extent of balsam fir, and other understory fuels, is unknown in

northeastern Minnesota, because of the cost of collecting field data (Menning and Stephens 2007) and difficulty of mapping understory vegetation using traditional satellite imagery structure (Lefsky et al. 2002, Riaño et al. 2003).

A proposed solution to mapping understory vegetation is the use of light detection and ranging technology, commonly known as lidar. Modern lidar has the ability to record multiple returns and can be used to penetrate the overstory canopy and measure understory vegetation components. Airborne lidar has been successfully used to measure vegetation structure from the lower levels of the canopy (Kramer et al. 2014, Kramer et al. 2016, Skowronski et al. 2007), where balsam fir typically grows. Airborne lidar data are available for all of Superior National Forest through the Minnesota Department of Natural Resources and were collected in spring of 2011 (Minnesota Department of Natural Resources 2014).

Objectives

The objectives of this investigation were to: 1) calculate a series of lidar metrics from the original point cloud; 2) determine if any one metric, or group of metrics, corresponds to field measurements of understory vegetation taken in the study area in Superior National Forest; 3) determine if there is anything in the lidar data that may be used to extract the balsam fir component of the understory vegetation; and 4) to map the understory vegetation using the lidar data.

Study Area

Superior National Forest is located in the arrowhead region of northeastern Minnesota (Figure 1). Superior National Forest is over 1.2 million hectares (3 million acres) in size, is home to the over 400,000 hectares (1 million acres) large Boundary Waters Canoe Area Wilderness (Figure 2), and is the eighth most visited national forest in the United States (United States Department of Agriculture n.d.).

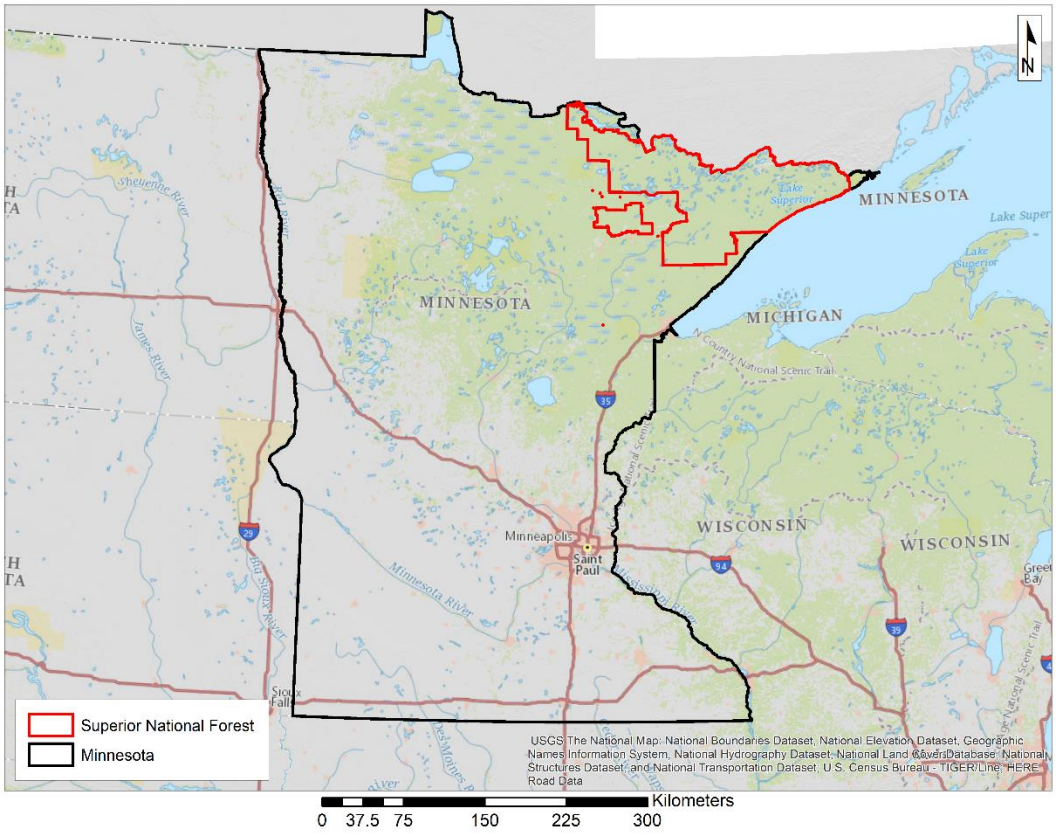

Figure 1. Map showing the location of Superior National Forest.

Superior National Forest is subject to cold, lengthy winters and brief summers, with average lows in the Boundary Waters ranging from -23˚ C (-10˚ F) in January to 15.5˚ C (60˚ F) in July and average highs ranging from -8˚ C (17˚ F) in January to 24˚ C

(75˚ F) in July. Average precipitation in the Boundary Waters region is around 686 mm (27”), with nearly 41 percent arriving in June, July, and August. Snowfall totals in the region typically reach up to 1778 mm (70”) (Heinselman 1973, Heinselman 1996).

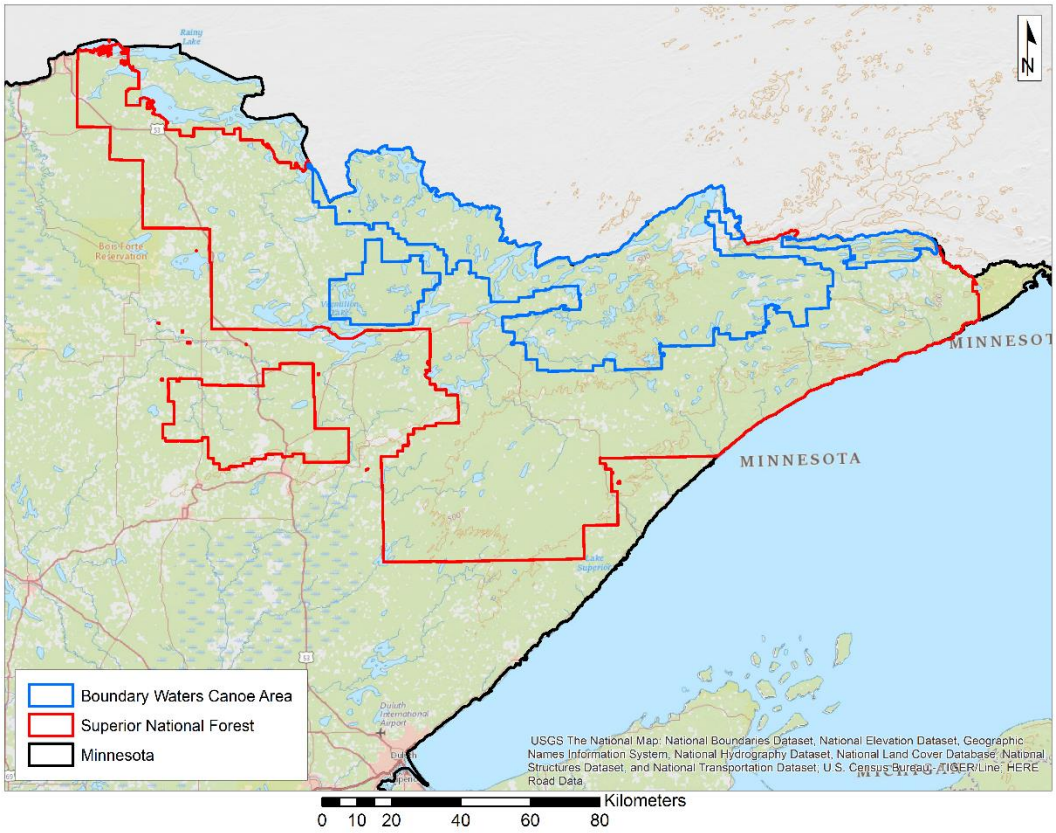

Figure 2. Superior National Forest and the Boundary Waters Canoe Area Wilderness.

Over 2,000 lakes are located within the boundary of the Superior National Forest, along with 5,470 km (3,400 miles) of streams (United States Department of Agriculture n.d.). The current topography of Superior National Forest, which has very little relief, was primarily shaped by the Wisconsin Glaciation, a period mainly of erosion in the area, with little deposition (Heinselman 1973, Heinselman 1996).

Superior National Forest is in the transition zone from the Great Lakes-St. Lawrence Forest Region to the Boreal Forest Regions and contains forest communities associated with both forest regions (Heinselman 1973, Heinselman 1996). Heinselman (1996) noted that, "An idealized view of the original north woods pictures unbroken forests of giant white and red pine, mixed here and there with spruce, balsam, jack pine, aspen, and birch," but, in reality, only about 25 percent of northeastern Minnesota contained mature red and white pine forest when Europeans began to settle, based on U.S. General Land Office Survey records and other accurate historical data sources. Plant communities in the region are broken into two types, upland types and lowland types. These plant communities werre shaped by natural disturbances, especially fire (Heinselman 1996).

The study area (Figure 3) for this research, which was established by the U.S. Forest Service, was a 35,613 hectare (88,000 acre) rectangular portion of the Superior National Forest and Boundary Waters Canoe Area Wilderness. The study area is located along the Gunflint Trail, with the southeast corner at the intersection of Greenwood Lake Road and Assinika Creek and the northwest at South Lake.

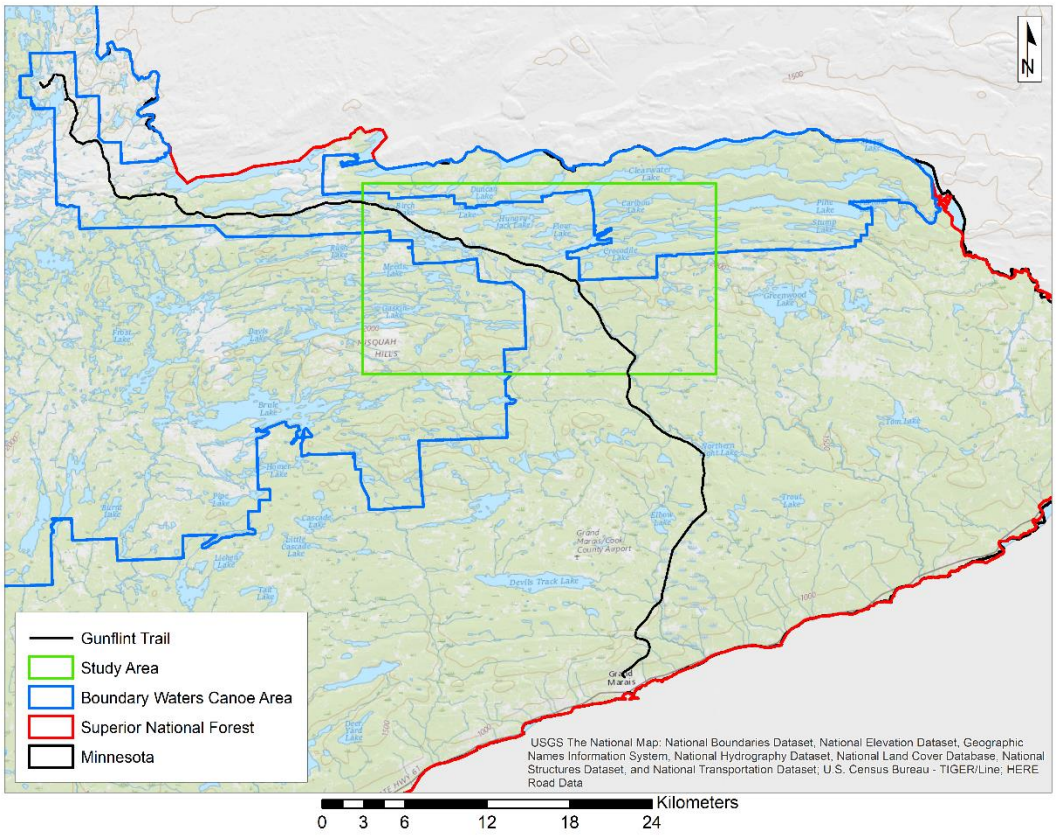

Figure 3. The location of the study area within Superior National Forest.

LITERATURE REVIEW

A Brief History of the Superior National Forest Region

The vegetation that is present in Superior National Forest today was very much affected by the glacial history of the region. Over the last 2 million years, numerous continental glaciers have traversed the region from the north, originating from the Hudson Bay-Labrador region. The Wisconsin Glaciation eliminated nearly all indication of the previous glaciations. There was little glacial deposition within the region; it was mainly an area of erosion. What little glacial drift that was deposited served as the parent material for the area's shallow, rock infused soils (Heinselman 1973, Heinselman 1996).

Tundra like vegetation appeared about 15,000 to 11,000 years ago. A dwarf birch shrubland arose around 10,300 years ago, about the same time the continental ice sheets started to retreat in the area. Spruce started growing in the shrubland around 10,000 years ago. The spruce forests reached their peak between 10,000 and 9,500 years ago, then jack pine began invading and the spruce were waning by 9,300 years ago. Other species present at that time included tamarack, northern white cedar, aspen, black ash, elm, oak, balsam fir, and paper birch. During the postglacial warm-dry period, about 8,000 to 4,000 years ago, the only major changes were the arrival of white and red pine in the region. Around 3,000 years ago, the pine species, especially white and red, began to decline and the boreal species, spruce, white cedar, and fir, were on the rise (Heinselman 1973, Heinselman 1996, Swain 1973).

The first evidence of human occupation in the area comes after glacial retreat from the region. Sometime between 10,000 and 7,000 years ago, the Plano people were living in what is now Superior National Forest. The indigenous people of the area were

using copper technology about 5,000 years ago and the Old Copper people were thought to have started making dugout canoes from red and white pine. That was likely to have been the origin of the canoe routes in the Boundary Waters Canoe Area. The birch-bark canoe was developed between 1,000 and 500 years ago, in the Woodland Tradition way of life (Heinselman 1996).

European explorers arrived in the Lake Superior region in the 1600's, although certain aspects of their culture and their tools had probably been known to the native populations earlier. With the arrival of the Europeans came the demand for fur and the start of the fur trade (Heinselman 1973, Heinselman 1996). Permanent European settlement came to the area around 1890, along with logging and mining. At the time of European arrival, the native peoples of the Boundary Waters region were existing within the natural system, but with the Europeans came more demands on the ecosystems (Heinselman 1996).

Iron ore was discovered in the Superior National Forest region in the last half of the 1800's (Heinselman 1973, Heinselman 1996). That discovery lead to the first of two major logging eras. The Big-Pine Logging Era lasted from 1895 to 1930 and was driven by the demand for white and red pine lumber for mine supports (Heinselman 1996). The first saw mill was built near Tower, Minnesota around 1890 (Heinselman 1973, Heinselman 1996).

The Pulpwood Logging Era started around 1935 and lasted until about 1978. The tree species that were in demand during the Pulpwood Logging Era were mainly jack pine and black spruce. The Pulpwood Logging Era had five driving factors, 1) the establishment of Superior National Forest in 1909 and U.S. Forest Service timber-

management program, 2) the maturation of large tracts of jack pine and black spruce on those newly created public lands, 3) the growth of the wood pulp and paper industries in Minnesota and neighboring states, 4) the development of mechanized logging and road construction, and 5) the construction of a road system that allowed for transportation of timber directly from Superior National Forest to the pulp mills (Heinselman 1973, Heinselman 1996).

Construction of gravel roads transformed the region's topography and changed the region's drainage networks(Heinselman 1996). The Big-Pine Logging Era transformed numerous red and white pine stands to aspen and birch stands, mainly through slash fires that killed any white and red pine seeds and remaining seedlings that were left following logging (Heinselman 1973, Heinselman 1996). Similarly, many of the jack pine and black spruce stands were transformed into red pine and white spruce or to aspen, birch, and fir stands by leaving those species uncut during the Pulpwood Logging Era. Logging converted large swaths of varying forest age classes to young forest classes with a more homogenous and reduced pine composition (Heinselman 1996).

Superior National Forest Ecosystems

The Boundary Waters region is a transition area. The Boreal Forest Region transitions to the Great Lakes-St. Lawrence Forest Region and hence, the Boundary Waters area contains constituents from each, though, the area does tend to be more Boreal in character. Boreal forest species are more adapted to cold climates than Great Lakes-St. Lawrence species. Boreal forests tend to be dominated by conifers, whereas

Great Lakes-St. Lawrence forests tend to be dominated by deciduous species. The boreal tree species that are present in the Superior National Forest region include jack pine, black spruce, white spruce, balsam fir, tamarack, northern white cedar, quaking aspen, and paper birch. The "idealized" red and white pines are Great Lakes-St. Lawrence Forest tree species, as are red maple, American elm, black ash, bigtooth aspen, balsam poplar, and the seldom seen yellow birch and basswood trees. Sugar maple, beech, and eastern hemlock, along with yellow birch, are the main dominant species of the Great Lakes-St. Lawrence forests, but are generally not found within the Boundary Waters. (Heinselman 1973, Heinselman 1996).

Plant communities within the Boundary Waters Canoe Area Wilderness region are broken into two types, upland types and lowland types. These communities are mainly defined by the dominant species found within them. There are thirteen types of upland plant communities: lichen, jack pine-oak, red pine, jack pine-black spruce, jack pine-fir, black spruce-feathermoss, maple-oak, aspen-birch, aspen-birch-white pine, maple-aspen-birch, maple-aspen-birch-fir, fir-birch, white cedar, and unclassified logged upland. There are eight different plant communities that comprise the lowland types. They are mixed conifer swamp forest, black spruce bog forest, tamarack bog forest, sphagnum-black spruce bog, ash-elm-swamp forest, alder-willow wetland, marsh and open muskeg, and open water communities (Heinselman 1996).

Vegetation communities are not static and usually develop after a disturbance or a combination of disturbances. There are nine major types of disturbance that affect the plant communities in the Boundary Waters region, they include, "windstorms, drought, flooding, fire, insects, mammals, plant diseases, humans (especially logging and

postlogging forestry activities), and succession (gradual natural changes in community composition and structure related to the passage of time, in the absence of major disturbances)." These major factors do not include the effects of climate change and plant migration (Heinselman 1996).

During the first half of the twentieth, ecology was dominated by the idea of succession until a "climax" system was reached. A climax ecosystem is one in which a "long-term equilibrium" exists between the vegetation and a geomorphically and climatically stable physical environment (Wright and Heinselman 2014).

Buell and Niering (1957) found the fir-spruce-birch forest community to be the climax forest type for northern Minnesota as they determined that in "five relatively undisturbed upland conifer forest stands… balsam fir, white spruce and paper birch were consistently present among the dominant trees." Buell and Niering (1957) also found that in all of the fir-spruce-birch forest stands, the stands were populated by pine species in the past, further supporting their ideas of successional climax. Balsam fir was found to have the highest percentages of both cover and basal area. Therefore, it was concluded that, "Fir reproduction is the most successful and will continue to maintain for the fir a dominant role," (Buell and Niering 1957).

Many investigators argued that the idea of a climax ecosystem did not fit everywhere (see Wright and Heinselman 2014) and that in many forest ecosystems, including the Great Lakes-St. Lawrence Forest and boreal forests, fire plays a critical role in ecosystem stability (Heinselman 1973, Heinselman 1996, Wright and Heinselman 2014). Wright and Heinselman (2014) thought that ecologists, "saw fire primarily as the

destroyer of the forest," a thing that was separate from the system that disrupted succession.

Wright and Heinselman (2014) thought that previous work on ecological succession lacked a proper historical context and that to understand fire dependent systems, viewpoints from time periods covering both a, "few hundred years" and, "many thousands of years" were necessary. Wright and Heinselman (2014) expressed this by stating:

> In a fire-dependent forest one may consider equilibria on two scales—a long-range equilibrium represented by the entire ecosystem in which fire is the principle environmental factor, and a short-range equilibrium, or climax in Clements' sense, which is occasionally approached in local elements of the mosaic that have escaped fire for a long time, such as an island in a lake. In a fire-dependent forest, the Clementsian climax is rarely reached, although continuation of the present policy of fire suppression would make its attainment more likely.

To obtain these two "perspectives," a reconstruction of the short-term fire history by coring trees and dating fire scars could be conducted. To view the long term ecological history, analyzing pollen and charcoal deposits in varved lake sediments would be necessary (Wright and Heinselman 2014).

Fire and Fuels in the Superior National Forest Region

Fire is not new to the Superior National Forest. Charcoal and pollen analyses from Lake of the Clouds sediment varves suggest that, "fire has been an important factor in determining the composition of the forest vegetation in the BWCA [Boundary Waters Canoe Area] during the past 10,000 yr at least." The analyses showed little change in fire or vegetation over the past 1,000 years. A "conservative" average rate of fire occurrence over that same interval was determined to be 70 to 80 years around Lake of the Clouds

(Swain 1973). Swain (1973) stated, "These periodic fires, which produce only short-term changes in the vegetation for 20-30 yr, have had a positive influence on the long-term stability."

Heinselman (1973) reconstructed and mapped fire histories within the virgin forests of the Boundary Waters Canoe Area by counting tree rings from fire scarred trees, using documentation from the U.S. General Land Office Survey township reports, using an incremental borer to determine stand ages from trees that reproduce well after fires, and using air photos (primarily from 1948) to aid in mapping forest stands and fire boundaries. Historical fires were mapped back until about 1600. The research indicates that fires frequently occurred in the area at short intervals, every 1 to 8 years. There were also numerous "major fire years," often accompanying periods of draught, which occurred at longer intervals. These major fire years were responsible much of the total burned area during the period of analysis (Heinselman 1973). Although any evidence would have been removed, Heinselman (1996) also speculated that many fires occurred between 1600 and 1895 in the areas that were logged for big pine.

Through his extensive research in the Boundary Waters, Heinselman (1973) postulated the idea of a "Natural Fire Rotation," which is "the average number of years required in nature to burn-over and reproduce an area equal to the total area under consideration." The Natural Fire Rotation for the Boundary Waters Canoe Area was determined to be around 100 years, with variation in stand types. Jack pine-black spruce stands and aspen-birch stands would typically have a shorter rotation. Whereas stands dominated by red and white pine would have a longer rotation, as would many lowland bog and swamp communities (Heinselman 1973, Heinselman 1996).

There are really 3 different historical fire regimes in the region. These fire regimes are related to how the dominate species are adapted to fire. The first fire regime is that of the red and white pine forests and presumably would apply to about 25 percent of the area (Heinselman 1996). Red and white pines are from the Great Lakes-St. Lawrence forest and are adapted to survive light to moderate intensity surface fires. Their cones are not serotinous and seed crops do not occur regularly (Heinselman 1973, Heinselman 1996). In this fire regime, these low intensity surface fires occur at return intervals of 10 to 50 years, with higher severity fires, which resulted in partial to total stand replacement, occurring at intervals of 150 to 200 years (Heinselman 1996).

The second fire regime is the dominant fire regime of the area and would cover around 67 percent of it. This fire regime is that of the boreal conifer and deciduous forest types, those dominated by jack pine, black spruce, quaking aspen, paper birch, or some combination thereof (Heinselman 1996). Jack pine has serotinous cones and black spruce has semiserotinous cones, making both dependent on fire to regenerate. Both aspen and birch spread seeds and are capable of sprouting from roots. This fire regime is characterized by high intensity surface fires and/or high severity crown fires (Heinselman 1973, Heinselman 1996). "Crowning occurs only in forests with abundant conifers, especially long-crowned conifers such as balsam fir, black spruce, white spruce, or cedar with live branches near the ground even in the shade of an overstory jack pine or red and white pine. These conifers can serve as fire ladders to help loft the flames up into the crowns of the overstory," (Heinselman 1996). Large portions of forest stands are killed in this fire regime and historically large areas would experience these types of fires in the region.

The third fire regime is less understood. It occurs in the low-lying wetland areas of the region. Evidence suggests that these swamp and bog forest communities experience, "regimes of long return-interval but fairly high-intensity surface fires or partial crown fires," (Heinselman 1996).

Heinselman (1996), through his decades of work in the area, viewed the Boundary Waters as, "a patchwork of various stand-age classes and forest communities resulting from a long history of fires of various sizes, shapes, and intensities." Heinselman (1973) states:

> My fire history studies conducted from 1966 to 1972, reported here, and the related researches of Swain (1972, this symposium), leave no doubt that fire was a dominant factor in the entire primeval system. Fire largely determined the species composition, age structure, and mosaic of successional stages of the forests, and thus also the habitat patterns for wildlife. It was also a major factor controlling nutrient cycles and energy pathways, and in maintaining the diversity, productivity, and stability of the whole ecosystem.

The idea that a climax forest, in the Superior National Forest region, was stable was rejected (Heinselman 1973, Heinselman 1996, Swain 1973, Wright and Heinselman 2014). The data of Heinselman (1973) show that there were a great number of balsam fir trees in his study areas that had not burned for a long time. He argued that a climax vegetation of fir-spruce-cedar-birch forest would not be more stable than "the heterogeneous vegetation mosaic produced by periodic fires." The two main arguments of Heinselman (1973) were that stands of fir and white spruce are "highly susceptible to the spruce budworm," and that mature fir and spruce can suffer from wind breakage and uprooting.

The expansion of balsam fir on the landscape is not a product of succession, as Buell and Niering (1957) envisioned it, but instead it is a primarily a result of man's

practices, especially fire suppression (Heinselman 1973, Heinselman 1996). Heinselman (1973) stated that, "what appears to be succession is really just suppression."

Superior National Forest was established by President Roosevelt in 1909 and with its establishment came fire suppression. The last year of major fires (prior to 1972), with the exception of the fires started during the severe drought in 1936 and the Little Sioux Burn in 1971, was 1910. As a consequence, the Natural Fire Rotation time rose from around 100 years pre-fire suppression to about 2,000 years for the period of 1911 to 1972 (Heinselman 1973, Heinselman 1996).

Fire exclusion is not the only culprit in balsam fir expansion, however, logging practices also share a role in converting large amounts of forest from red and white pine and jack pine-black spruce forest types to aspen-birch-fir compositions. During the Big-Pine Logging Era, numerous red and white pine stands were converted to aspen and birch as a result of slash fires. These slash fires killed any red and white pine seeds, seedlings, and seed trees that remained after harvesting was completed. The Pulpwood Logging Era focused mainly on harvesting jack pine and black spruce and took place between the years of 1935 to 1978 (fire suppression). When the jack pine and black spruce were logged, any aspen, birch, and fir trees that were present were often left untouched, and since fire was suppressed, the aspen, birch, and fir trees were allowed to expand on the landscape, while the serotinous and semiserotinous seed cones for the jack pine and black spruce rotted away (Heinselman 1996). Heinselman (1996) calculated that vegetation shifts brought on by logging are orders of magnitude greater than those brought about by climatic changes during the last 3,000 years for the Boundary Waters Canoe Area.

Another hypothesized cause for the spread of balsam fir is climate change (Heinselman 1996). While it has been shown that increases in temperature and carbon dioxide increase photosynthetic rates of forest species over short time frames (Mooney et al. 1991), it is unlikely that climate change would favor balsam fir, as balsam fir is a boreal species and the boreal forest is already at its southern limits in northeastern Minnesota (Heinselman 1996).

Balsam fir expansion is not solely limited to Superior National Forest, it has also been occurring in eastern Canada (Blais 1983a, Blais 1983b). Fire suppression started around 1920 in eastern Canada, decreasing the amount of forest area that burned and encouraging succession to fir—spruce stands. Logging has also produced similar patterns in eastern Canada, where large scale harvesting of white pine stands resulted in stand replacement by fir and spruce. Stands logged for pulpwood also increased the amount of fir on the landscape in eastern Canada (Blais 1983b). Increases in northern white cedar, balsam fir, and spruce and decreases in pine species have also been noted in national wildlife refuges in northern Minnesota and Michigan (Corace III et al. 2012).

Other researchers have investigated the changing landscapes of northern Minnesota. Frelich and Reich (1998) conceptualized three different models of succession in the boreal forests of northern Minnesota based upon the severity of disturbance. They noted that stand composition has changed because of fire exclusion and timber harvesting primarily, but that other less severe disturbances also transpire, such as wind blowdown and insect infestations (Frelich and Reich 1998).

Friedman, Reich, and Frelich (2001) used General Land Office survey data to examine species composition and distribution for the presettlement period in the

Arrowhead Region of Minnesota. When they looked at species composition, they found that black spruce accounted for 20.7%; paper birch, 15.1%; larch, 15.0%; aspen 10.8%, balsam fir, 9.4%; jack pine, 7.8%; white pine, 6.3%; northern white cedar, 6.1%; red pine, 2.7%; and other species accounted for the remaining 6.1% of the composition. Friedman, Reich, and Frelich (2001) also calculated the proportion of the basal area that each species accounted for: white pine 20.1%; paper birch 14.0%; black spruce 13.4%, larch 11.0%, aspen 7.7%, red pine 7.3%, northern white cedar 6.0%; jack pine 5.7%; balsam fir 5.2%; and other species 9.6% of the total basal area.

Expanding on the work of Friedman, Reich, and Frelich (2001), Friedman and Reich (2005) compared circa 1880 General Land Office survey data to U.S. Forest Service Forest Inventory and Analysis data from 1990. The data were mapped in 10 km by 10 km grid cells. Protected areas, such as the Boundary Waters Canoe Area Wilderness, were excluded from analysis because those types of areas are not inventoried by the Forest Inventory and Analysis. The authors found that "Evidence developed in this analysis indicates that the forest has undergone substantial compositional and spatial change." Some of the most dramatic increases in species relative abundance across the region included increases in aspen from 10.7% to 29.9%, balsam fir from 9.5% to 15.7%, maple from 1.4% to 6.3%, and ash from 1.4% to 5.0%. Many tree species declined in abundance. Spruce (both white and black) dropped from 21.0% to 16.2%, larch decreased from 15.4% to 2.9%, paper birch was reduced from 14.9% to 11.3%, jack pine from 6.5% to 2.4%, white pine went from 6.2% to 1.1%, and red pine decreased from 2.8% to 1.9%. Friedman and Reich (2005) report that not all pine trees were identified by species in the GLO data and that all in all, pines decreased from 17.0% to 5.4% from

the 1880s to 1990. In terms of proportional basal area, aspen had the largest increase (7.6% to 34.8%) and white pine had the largest decrease (20.4% to 5.2%) and all pines decreased from 36.2% to 12.5% in relative basal area.

Friedman and Reich (2005) stated:

> The substitution of logging for fire as the dominant disturbance regime in northeastern Minnesota led to major alterations of composition and dominance in this region, regardless of whether measured across physiographic zones, associations with soils, by species, or by community type… Community type conversions, as we observed them, represent a replacement of the presettlement forest matrix with a new and substantially more homogeneous matrix (i.e., the dominance of aspen is much greater than the dominance of any species in the presettlement time).

Additionally, researchers have tried to peer into the future to see how the ecosystems in northeastern Minnesota might continue to change. Scheller et al. (2005) used a model called LANDIS to predict the structure and composition of the Boundary Waters Canoe Area after 300 years. With a starting year of 1995, the researchers ran four different scenarios, a fire rotation period of 50 years, a fire rotation period of 100 years, a fire rotation period of 300 years, and no fire. The fire rotation periods of 50 and 100 years were thought to best express the historic fire rotation as described by Heinselman (1973) and others (Scheller et al. 2005). Scheller et al. (2005) concluded that after 300 years, species composition and structure would approach historic levels with a return of fire rotation periods of 50 or 100 years, with 50 years being the best as it would reduce high severity fires over time. The authors also determined that a fire rotation period of 300 years would preserve all historic species, but would not produce a comprehensive landscape restoration. Scheller et al. (2005) determined that no fire was the worst option for the Boundary Waters and that excluding fire would result in a landscape dominated

by white spruce and balsam fir, with species like jack pine, white pine, and red pine being removed from the forest over time.

Prescribed fire and prescribed natural fire (fire initiated by lightning which is allowed to burn under certain conditions) were introduced as management tools in Superior National Forest in 1987 (Heinselman 1996). Other means of reducing fuels have also been explored.

Gilmore et al. (2003) conducted experiments to compare fuel reduction treatments in an area of Superior National Forest that experienced a major blowdown in 1999. They worked in two forest types, jack pine and aspen/birch. It was determined prescribed burning did an excellent job at decreasing fine fuels, but was less effective than other treatments at decreasing heavy fuels, whereas treatments of salvage logging and machine piling worked to reduce both fine and heavy fuel loads (Gilmore et al. 2003).

The Ham Lake fire of 2007 occurred in the blowdown area in which Gilmore et al. (2003) conducted their research. Fraver et al. (2011) looked at the impacts that salvage logging had on reducing fire severity and fire intensity across three conditions in the area: blowdown, salvage, and fire; blowdown and fire; and fire only. It was concluded that salvage logging reduced, "the intensity (heat released) of the subsequent fire," and tree-crown fire severity, but increased forest-floor fire severity (Fraver et al. 2011).

Studies have also examined how harvesting forest fuels for biomass might offset the cost of wildfire fuel reduction treatments in Superior National Forest. Demchik et al. (2009) assessed the practicality of selling biomass from fuel reduction treatments in three wildland-urban interface areas in Superior National Forest. It was determined that sale

revenues did not cover harvesting costs, but could be used to offset costs in areas less than 160 km (100 miles) from a biomass plant (Demchik et al. 2009). Abbas et al. (2011) conducted similar research in Lutsen Township, Minnesota. The authors compared the costs of conventional fuel reduction techniques (prescribed burning, mastication, harvesting, and piling and burning) versus using the wood products as a source for bioenergy and concluded that, conventional fuel treatments costs were lower than harvesting and selling fuels for biomass (Abbas et al. 2011).

Balsam Fir Characteristics

Balsam fir is a conifer that is spread across much of north central to northeastern North America. The range of balsam fir runs from northwestern Alberta through central Saskatchewan and Manitoba, south to northern Minnesota east to New England, and back north to Newfoundland. Balsam fir has low economic value, but is most often harvested for pulpwood, light weight lumber, and Christmas trees (Uchytil 1991).

Balsam fir can grow on many soils, but mainly grows on acidic Spodosols, Histisols, and Inceptisols, and can be found on both upland and lowland sites. Balsam fir trees produce many seeds after the trees reach the age of 20, with regular seed generation at the age of 30. Balsam fir seeds are winged and wind scatters the seeds short distances (25 m – 60 m or 80’ – 200’) from the source tree. Seed productivity is low and good seeds can only take root within one year. Balsam fir trees have continuous canopies that reach the surface. When many balsam fir trees are clustered together, the lower limbs are usually lifeless, but remain attached to the stems (Uchytil 1991).

Balsam fir is considered a late successional species. It is rarely found in pure stands and is rarely the dominant species. Because seedlings are shade tolerant and can grow in many site conditions, balsam fir can develop in the understory of pines, aspen, and birch overstory trees of boreal forests (Figure 4). Areas with high proportions of balsam fir are vulnerable to spruce budworm infestations, especially if many of the balsam fir trees have reached the age of 50 (Uchytil 1991).

Figure 4. Balsam fir growing in the understory of a red and white pine stand in the George Washington Pines area of Superior National Forest. Photo taken by Birgit Peterson, USGS EROS, in May 2017.

Much research has been conducted on spruce budworm outbreaks. Blais (1983b) analyzed the frequency of spruce budworm outbreaks over the past 3 centuries for eastern Canada. The results indicated that not only are the outbreaks occurring at shorter

intervals, but they are more extensive, and more severe when comparing the 20th century data to the previous centuries. This was primarily attributed to "man's actions," through activities that promote balsam fir (the principle host species) such as logging practices, fire control, and insecticides (Blais 1983b). Blais (1983b) does note that, "Spruce budworm outbreaks are a natural phenomenon; the result of the coevolution of a plant—insect system. As such they are ultimately beneficial to the forest. By destroying considerable quantities of mature and overmature fir they prevent the perpetuation of decadent stands and bring about a rejuvenation of the forest."

Stocks (1987) conducted a series of experimental fires in central Ontario, Canada. Balsam fir was the main tree species within the plots. The sparse overstory contained white pine, jack pine, white spruce, and birch. Spruce budworm had infected the balsam fir for at least five years and all balsam fir were dead in the test plots and top breakage of the trees was occurring. Stocks (1987) found that spring fires behaved explosively, noting that all the successful experimental fires, "reached intensity levels well beyond the control capability of any forest fire agency." The research also indicated that green shrubs inhibited summer fires in the initial years following defoliation and tree death of balsam fir, but that, "After 4 -5 years, however, enough dead and down surface fuel accumulates through balsam fir decomposition to offset this dampening effect of understory vegetation, and summer fires will spread through the fuel complex," (Stocks 1987). Stocks (1987) determined that forest fire potential is at its highest 5 – 8 years after balsam fir death, because of the high surface fuel loads that result from crown breakage, windthrow, and the arrangement of the fuel complex that results from those two

processes. After that peak period of high fire potential, decomposition causes the risk to decrease (Stocks 1987).

Fleming, Candau, and McAlpine (2002) examined Canada's Forest Insect and Disease Survey records on spruce budworm outbreaks and compared them to large wildfire occurrences from 1941 to 1996. The authors found was that wildfire followed spruce budworm outbreaks most frequently 3 – 9 years after the outbreak, with some geographic variability from climatic differences. Fire suppression was also found to be an important factor in eastern Canada. Fleming, Candau, and McAlpine (2002) indicated that climate change resulting in a warmer and drier regional climate could cause a positive feedback loop in which more warming causes more spruce budworm outbreaks and mores fires, which releases more carbon, which leads to more warming, and so on, and so on.

Sturtevant et al. (2012) used a landscape succession and disturbance model, LANDIS-II, to study the interactions between spruce budworms and wildfire in the Boundary Waters Canoe Area. The authors argued against the idea that insect disturbances increase the risk of wildfires and concluded that spruce budworm outbreaks actually reduce wildfire risk over the time scales of decades to centuries. "Our study suggests that budworm serves as a natural thinning agent that decreases live ladder fuels by periodically reducing balsam fir content," (Sturtevant et al. 2012). This is not in complete disagreement with previous research, as Stocks (1987) did note that fire risk decreases following the peak 5 – 8 year period of highest risk.

Wolter et al. (2008) used Landsat data to map relative basal areas of spruce budworm host species (balsam fir and white spruce) over an area that included all of

Superior National Forest, Voyagers National Park, and Quetico Provincial Park (Canada). The authors also developed models for relative basal area of total deciduous species and total coniferous species. Wolter, Townsend, and Sturtevant (2009) expanded on this work by using SPOT-5 data to estimate tree canopy diameter, bole diameter at breast height, tree height, crown closure, vertical length of live crown, and basal area – forest structure metrics that can be used to determine susceptibility to pest infestation. The study area for this research was mainly in Lake County, Minnesota and included a portion of Superior National Forest. Wolter and Townsend (2011) further expanded upon the previous studies by combining synthetic aperture radar data from Radarsat-1 and PALSAR, with SPOT-5 and Landsat imagery for the same study area as was used by Wolter, Townsend, and Sturtevant (2009). Wolter and Townsend (2011) were able to map the relative basal areas of aspen, birch, black ash, maple, white pine, jack pine, red pine, eastern larch, black spruce, balsam fir, white spruce, and white cedar by combining the imagery. From their results, Wolter and Townsend (2011) also generated a forest cover type map. Wolter and Townsend's (2011) balsam fir map shows that it is, "ubiquitous on the landscape, but never at high relative BA (usually <10-20%). This is consistent with field observations showing the balsam fir rarely dominates the overstory in stands where it occurs."

As shown in the research discussed above, balsam fir is a serious fire risk for a large portion of North America. While long term fire risk may not be affected by spruce budworm outbreaks (Sturtevant et al. 2012), short term risk seems to increase (Fleming, Candau, and McAlpine 2002, Stocks 1987), and efforts have taken place to map areas

vulnerable to outbreaks (Wolter et al. 2008, Wolter, Townsend, and Sturtevant 2009, Wolter and Townsend 2011). So, the question is, how do fire and balsam fir interact?

"Balsam fir is easily killed by fire," this is mainly because of its, "thin, resinous, easily ignitable bark and shallow roots," (Uchytil 1991). Because of balsam fir's continuous crown, which is often full of dead lower limbs, and easily ignited needles and bark, it can act as a ladder fuel for fire, transferring fire from the surface into the upper canopy, leading to crown fires (Abbas et al. 2011, Heinselman 1973, Uchytil 1991). Even though post fire conditions are good for germination, balsam fir seeds are usually eliminated by fire. This results in stands that are free from balsam fir for 30 to 50 years after fire occurrences, unless isolated unburned areas occur inside the burn or adjacent to the burned area (Uchytil 1991).

Lidar

As discussed above, balsam fir grows in the understory and is a ladder fuel that can turn surface fires into crown fires. Traditional efforts to measure ladder fuels in the field can be time consuming and costly, or they can rely on unrepeatable, imprecise visual assessments. Additionally, it is difficult to extrapolate these efforts over large areas (Menning and Stephens 2007).

Menning and Stephens (2007) have developed a relatively rapid field method for assessing ladder fuel risk which they call "ladder fuel hazards." The approach allows field technicians to work through a flow chart and classify the ladder fuel hazard into one of five categories, ranging from no risk to high risk. The classification scheme is based on whether or not forest is present, the degree of low fuel clumping, and the vertical

continuity of the fuels (Menning and Stephens 2007). Unfortunately, this method is only semiquantitative and Menning and Stephens (2007) indicated that their method would allow a large amount of data to be collected over a vast area, but that there would be little data collected at any given site.

As mentioned by Menning and Stephens (2007), fuels that are present in the understory are difficult to map over large areas. Conventional passive remote sensing technologies also have difficulty measuring sub canopy forest structure (Lefsky et al. 2002, Riaño et al. 2003). Menning and Stephens (2007) mentioned lidar as a "promising" technology for making "three-dimensional models of forest vegetation."

Light Detection and Ranging, commonly known as lidar, is an active remote sensing technology that employs laser ranging to provide three-dimensional point information for objects, including Earth's surface, vegetation, and infrastructure (e.g. buildings, power lines, road surfaces, etc.) Lidar data are often acquired from the air and that process may be known as airborne laser scanning (Renslow 2012, Wehr and Lohr 1999). Lidar data come in the form of a three-dimensional point cloud that can be georeferenced. Each point in the cloud can also be classified (e.g. ground, water, noise) (Renslow 2012).

Mapping topography is probably the most well-developed use of lidar (Lefsky et al. 2002), where it has many uses, including determining areas likely to flood or which could be subjected landslides (Renslow 2012, Wehr and Lohr 1999). However, lidar also has the capability to measure vegetation structure in three dimensions, something that traditional passive remote sensing technologies have been unable to do, making lidar an

important tool in ecological surveys (Lefsky et al. 2002) and fuels measurements (Andersen, McGaughey, and Reutebuch 2005, Riaño et al. 2003).

Canopy height determination is one of the most basic and obvious uses of lidar, as is the measurement of canopy cover. These measurements started becoming common in the 1990's. The development of discreet return lidar systems (systems which can record multiple returns from each individual laser pulse) have allowed for more complex measurements to be made (Lefsky et al. 2002).

Naesset (2002) looked at using airborne lidar to estimate mean height, dominant height, mean diameter, stem number, basal area, and volume, so that stands might be inventoried in a Norway spruce and Scots pine forest in southeastern Norway. The lidar data contained first and last return points and were used to generate metrics that included distribution quantiles, mean values, maximum values, coefficients of variation, and canopy densities. In what the authors call a "two-stage procedure," these metrics were correlated with field data. Coefficients of determination ranged from 0.39 to 0.95 (Naesset 2002).

Riaño et al. (2003) used a small (2,000 m by 190 m) test site, in a spruce dominated forest outside of Ravensburg, Germany, and a last return airborne lidar system to determine fuel model variables. By normalizing the point cloud into a height above ground format (essentially making the ground a planer surface), the authors were able to use cluster analysis to separate overstory and understory components (Riaño et al. 2003).

Using an approach similar to that of Naesset (2002), Andersen, McGaughey, and Reutebuch (2005) collected field data for "representative" trees in a Douglas-fir and western hemlock forest in the state of Washington. These measurements were used to

estimate canopy fuel weight, canopy bulk density, canopy base height, and canopy height. Maximum height, mean height, and coefficient of variation of the lidar heights, the 20th, 50th, 75th, and 90th height percentiles, and a canopy density metric were calculated from an airborne lidar system that produced 4 returns. After developing the regression models, the canopy fuel metrics were mapped using GIS. The maps, "represent spatially-explicit data layers that can be used as direct inputs into fire behavior models to support the analysis of fire risk and the implementation of fuel mitigation programs," (Andersen, McGaughey, and Reutebuch 2005).

Skowronski et al. (2007) built an "off the shelf" lidar system that they felt would be affordable for wildfire management agencies. The lidar was flown from a helicopter and was used to estimate tree biomass and ladder fuel presence in the Pinelands of New Jersey, which contains pine, oak, and wetland forests. The lidar system recorded first returns and data was collected during spring leaf off conditions. Points were separated into bins, based upon the height above ground of the return. Good agreement was found between the percentages of returns in the 1 – 2 m and the 3 – 4 m height bins and the presence of understory fuels. It was concluded that, "aggregating data into height bins produces key information on the distribution of live fuels in the understory and sub-canopy," (Skowronski et al. 2007).

Goodwin et al. (2007) looked at the possibility of using a first return airborne lidar system to map understory vegetation in a dense coniferous forest on Vancouver Island, British Colombia. The two primary objectives of this research were to 1) determine if first return lidar could be used to measure understory cover in areas that were 30 X 30 m and 15 X 15 m and 2) see if lidar could detect changes in the percentage

of understory cover. Goodwin et al. (2007) found "strong positive relationships," between the field data they collected and the lidar, when they looked at returns greater than 0.5 m and less than or equal to 4.0 m in height. It was determined that lidar was good at mapping understory cover at the plot scale, but not as good at the sub-plot scale, this was attributed to the accuracy of the hand-held GPS unit that was used.

Wing et al. (2012) worked in the Blacks Mountain Experimental Forest, a ponderosa pine forest in northeastern California. A metric called "understory lidar cover density" was developed from field measurements of understory shrubs. Lidar points were filtered for this metric based upon the intensity of the return. Models created using weighted liner regression and beta regression produced R^2 values ranging from 0.7 – 0.8. Wing et al. (2012) found a problem similar to that of Goodwin et al. (2007) stating, "An important step in any airborne lidar data analysis for forestry applications is verification of geo-registered plot locations. Inaccurate plot locations can be one of the largest sources of model error found in many types of airborne lidar analysis." Wing et al. (2012) also concluded, "Although it is yet untested, we hypothesize that this method will not perform as well in forest types that contain an abundance of understory vegetation under dense overstory cover conditions, or where the understory and overstory layers intermix." Other authors have reported issues detecting understory vegetation in dense overstory canopies (García et al. 2011, Goodwin et al. 2007, Simonson, Allen, and Coomes 2014).

Kramer et al. (2014) used lidar data to determine if ladder fuels reduction treatments had been conducted in areas of the Meadow Valley of Plumas National Forest, northern Sierra Nevada of California. Fifty-three lidar metrics were calculated from the

original point cloud to compare to field data. Using classification tree analysis, (Kramer et al. 2014) determined that the percent cover between the heights of 2 – 4 m was the best lidar metric for distinguishing between treated and untreated areas. "While it was not our goal to create an absolute metric for ladder fuels, 2 – 4 m cover could be used as a proxy for the relative abundance of ladder fuel across a landscape," (Kramer et al. 2014).

Kramer et al. (2016) developed a new field approach to measure ladder fuels. The authors used a wildlife board based on those developed by Jones (1968) and Nudds (1977) to measure vertical fuel coverage. Using photographs taken at field plots, the authors visually assessed "ladder fuel cover," which is an estimate of the vertical cover in a plot, from the surface to 4 m in height. Averaged ladder fuel cover estimates were compared to several lidar metrics in the Klamath Mountains in northern California, using multiple linear regression. It was determined that the percentage of all lidar returns between 1 and 8 meters, the standard deviation of lidar point heights above 2 meters, and the percentage of first return lidar points between 8 and 16 meters produced the best model ($R^2 = 0.73$). Ladder fuel cover was mapped for the lidar coverage area (Kramer et al. 2016).

UNDERSTORY COVER MODEL DEVELOPMENT

Introduction

This chapter outlines the development of the major product of this thesis, the understory cover model. The literature review leads right into this chapter. Much of the development of the understory cover model is similar to the work conducted by Kramer et al. (2016).

Materials and Methods

Field Data Collection

In order to pursue the problem of mapping understory forest fire fuels in Superior National Forest, it was necessary to collect field data which could be used in conjunction with the airborne lidar data collected for the Minnesota Department of Natural Resources. Field data were collected using the cover line intercept method.

The line intercept method is a sampling method described in the U.S. Forest Service Fire Effects Monitoring and Inventory System, or FIREMON. "FIREMON consists of standardized sampling methods and manuals, field forms, database, analysis program, and an image analysis guide so that fire managers can 1) design a fire effects monitoring project, 2) collect and store sampled data, 3) statistically analyze and summarize the data, 4) link the data with satellite imagery, and 5) map the sampled data across the landscape using image processing," (Lutes et al. 2006). FIREMON's principal goal is "to measure the immediate and long-term effects of a planned or unplanned fire

on critical ecosystem characteristics so that fire managers can evaluate the impact of that fire on ecosystem health and integrity," (Lutes et al. 2006).

The line intercept method "is used to assess changes in plant species cover for a macroplot… Cover is recorded as the number of ft (m) intercepted by each species along a transect. Percent cover is calculated by dividing the number of ft (m) intercepted by each species by the total length of the transect," (Lutes et al. 2006). The line intercept method was developed by A. G. Tansley and T. F. Chipp in 1926 (see Tansley A. G., and T. F. Chipp. eds. 1926. *Aims and Methods in the Study of Vegetation*. London: British Empire Vegetation Committee 383 p.). The FIREMON line intercept sampling scheme is designed to be adaptable, so that procedures may be changed to make the fieldwork more efficient (Lutes et al. 2006).

To start the line intercept sampling, the predetermined plot center is located, then the tape measure is pulled in the desired azimuthal direction for the initial transect. The suggested transect length is 66 feet (20 meters) and suggested number of transects is 5. The transect starts at 0 feet, which is the plot center, and ends at the desired transect length, which would be 66 feet, if the recommend length is used. Starting from 0 feet and measuring out to the end, the horizontal linear length of each plant that intercepts the transect is measured to the nearest tenth of a foot. The species, size class (which helps to determine vertical distribution), and plant status (either live, dead, or NA) are recorded for each plant intercept. In the case of intermingling canopies, if the plants are of the same species, overlap is not differentiated, however, where overlap occurs between different species, it is recorded. Also, gaps of less than 2 inches (0.17 feet or 5.1 centimeters) are usually ignored. The height of each intercept, over the transect, is also

measured. The process is repeated for the additional transects. The percent cover is calculated by summing the intercepts for each individual species and dividing by the transect length (Figures 5 and 6) (Lutes et al. 2006).

Specific modifications from the recommended procedures for this project included: changing the number of transects used to either 1 or 4 transects per plot, measuring only trees, and not determining plant status (e.g. live or dead). The transects were measured as transect 1 at an azimuth of 0 degrees, or north; transect 2 at an azimuth of 90 degrees, or east; transect 3 at an azimuth of 180 degrees, or south; and transect 4 at an azimuth of 270 degrees, or west. In the case where only one transect was measured, it was transect 1.

Figure LI-1—The measuring tape is stretched taught below, in, or above the canopy vegetation, whichever position allows the easiest estimation of cover without the tape zigzagging around plants.

USDA Forest Service Gen. Tech. Rep. RMRS-GTR-164-CD. 2006

Figure 5. Visual representation of a line intercept transect (Lutes et al. 2006).

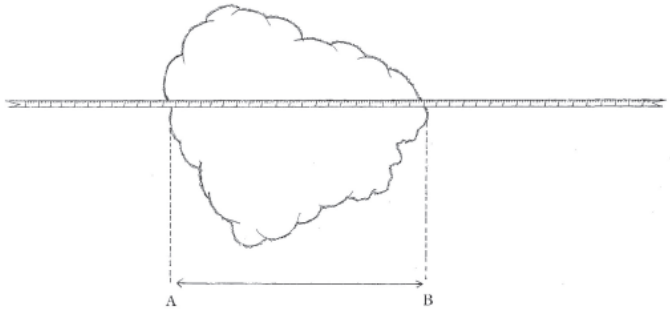

Figure LI-2—Measure cover intercept in feet (m) along the measuring tape. Since canopy intercept can vary on each side of the measuring tape, measure intercept on one side of the measuring tape only. We suggest using the right side as you move along the tape. Record the start of the plant intercept (A) in the Start field and the end intercept (B) in the Stop field.

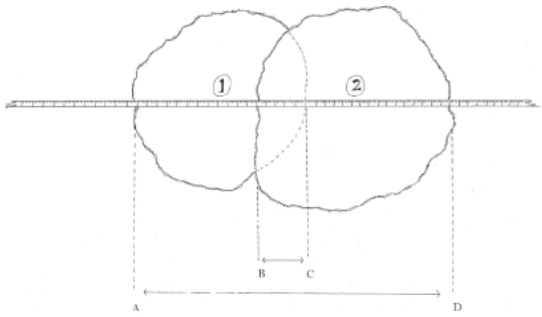

Figure LI-3—Canopy overlap (points B to C) is not measured if the canopy of two or more plants of the same species overlap. For example, if shrubs 1 and 2 are the same species, then the canopy intercept is measured from points A to D. If shrubs 1 and 2 are different species, then canopy intercept is measured from points A to C for shrub species 1 and from points B to D for shrub species 2.

USDA Forest Service Gen. Tech. Rep. RMRS-GTR-164-CD. 2006 LI-7

Figure 6. Diagrams demonstrating the process for measuring the horizontal distribution of the canopy (Lutes et al. 2006).

The U.S. Forest Service established the field plots for data collection (Figure 7). Field data were collected in two phases. During phase 1, cover line intercept data were collected for trees that fell into the pole tree class or larger (trees with a diameter at breast height of 12.5 centimeters or greater). During phase 2, cover line intercept data were collected for seedlings and saplings (trees with a diameter at breast height less than 12.5 centimeters). For the purposes of this research, the phase 2 data will be considered to represent the understory vegetation (Table 1).

In the field season of 2015, multiple U.S. Forest Service field crews collected phase 1 data at 130 plots and phase 2 data at 33 of those plots. In July of 2016, using the cover line intercept instructions provided to them by the U.S. Forest Service, a team from

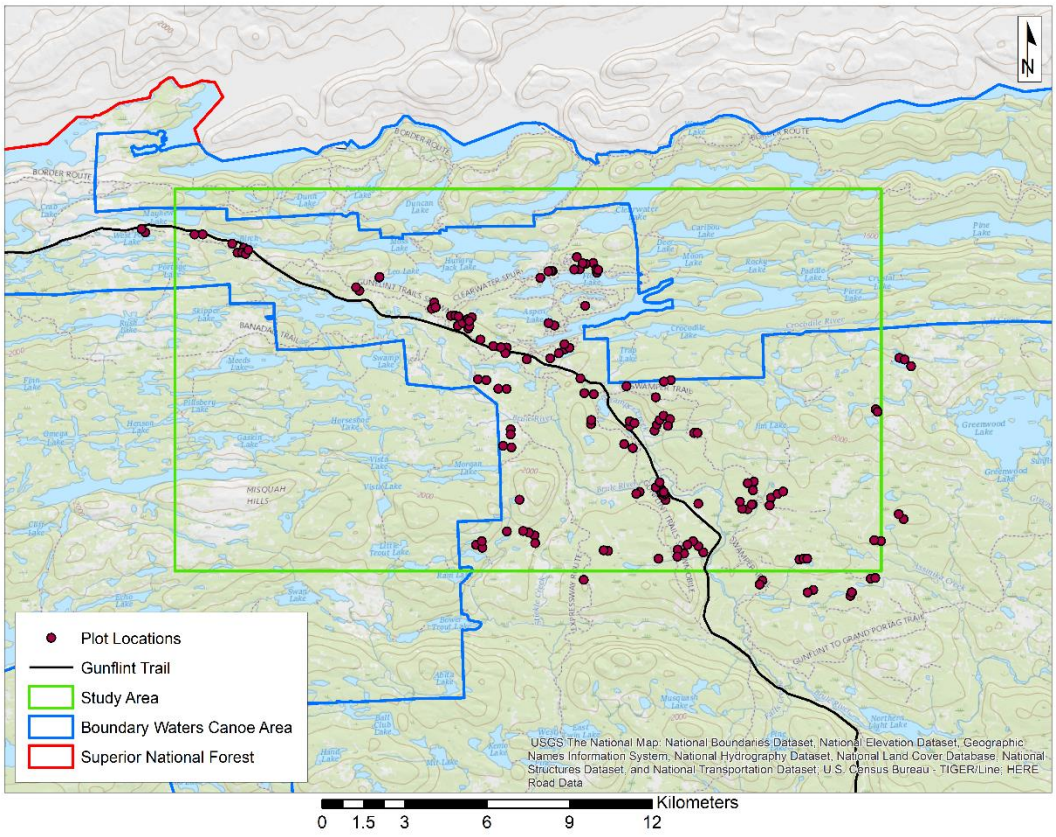

Figure 7. The locations of the field plots.

Name	Description
Very Large Trees	Trees ≥ 80 cm dbh
Large Trees	Trees 50 cm - < 80 cm dbh
Medium Trees	Trees 25 cm - < 50 cm dbh
Pole Trees	Trees 12.5 cm - < 25 cm dbh
Saplings	Trees 2.5 cm - < 12.5 cm dbh
Seedlings	Trees < 2.5 cm dbh or <1.5 m height

Table 1. The cover line intercept tree classification scheme, adapted from Lutes et al. (2006) (dbh = diameter at breast height).

the U.S. Geological Survey Earth Resources Observation and Science Center (commonly known as EROS) collected phase 2 data at an additional 13 plots (Figure 8). In addition to the initial field data collection efforts, the EROS team collected validation data at 24

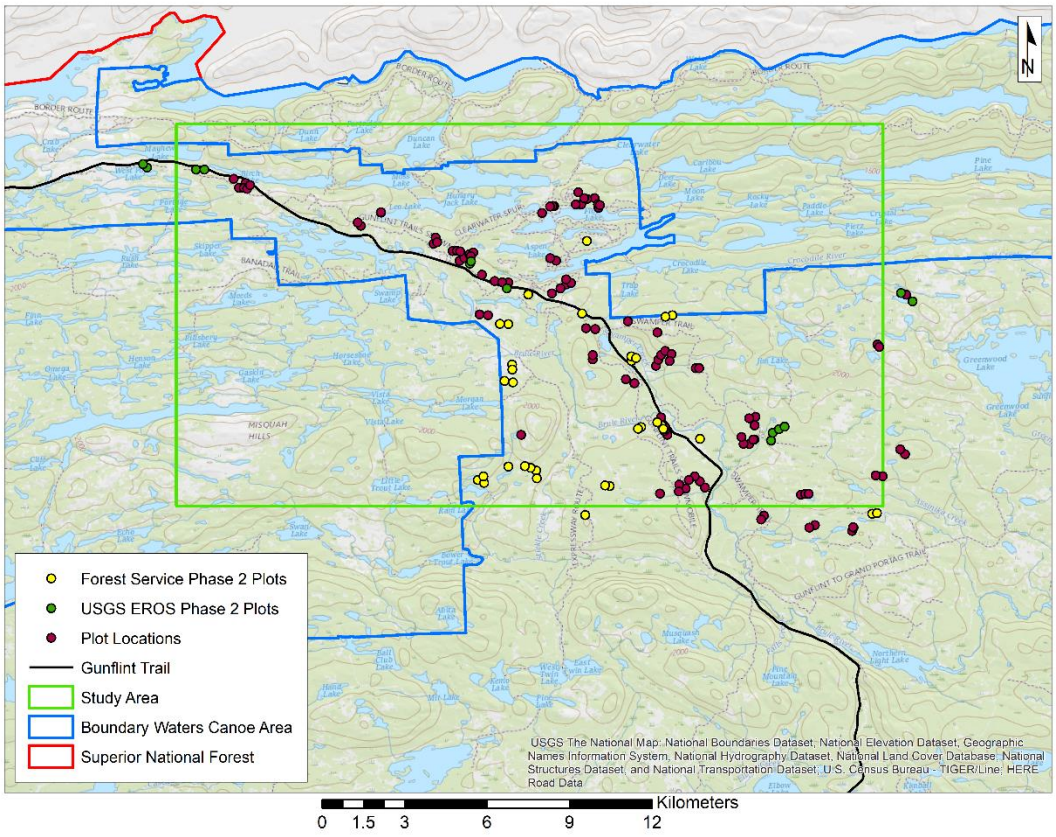

Figure 8. Phase 2 plot locations.

plots in May of 2017 (Figure 9). During both field excursions by the EROS team, photos were taken from the plot center looking out along each transect (Figures 10 and 11).

All of the field data were compiled into one format using Microsoft Excel (2016). The software was used to convert the cover line intercept measurements from the U.S. customary system into the metric system. For each plot, every transect that was measured had the length of the canopy that intercepted the transect calculated by species. Those individual intercepts were summed to obtain the total amount of transect canopy (note that as a consequence of the cover line intercept method measuring different tree

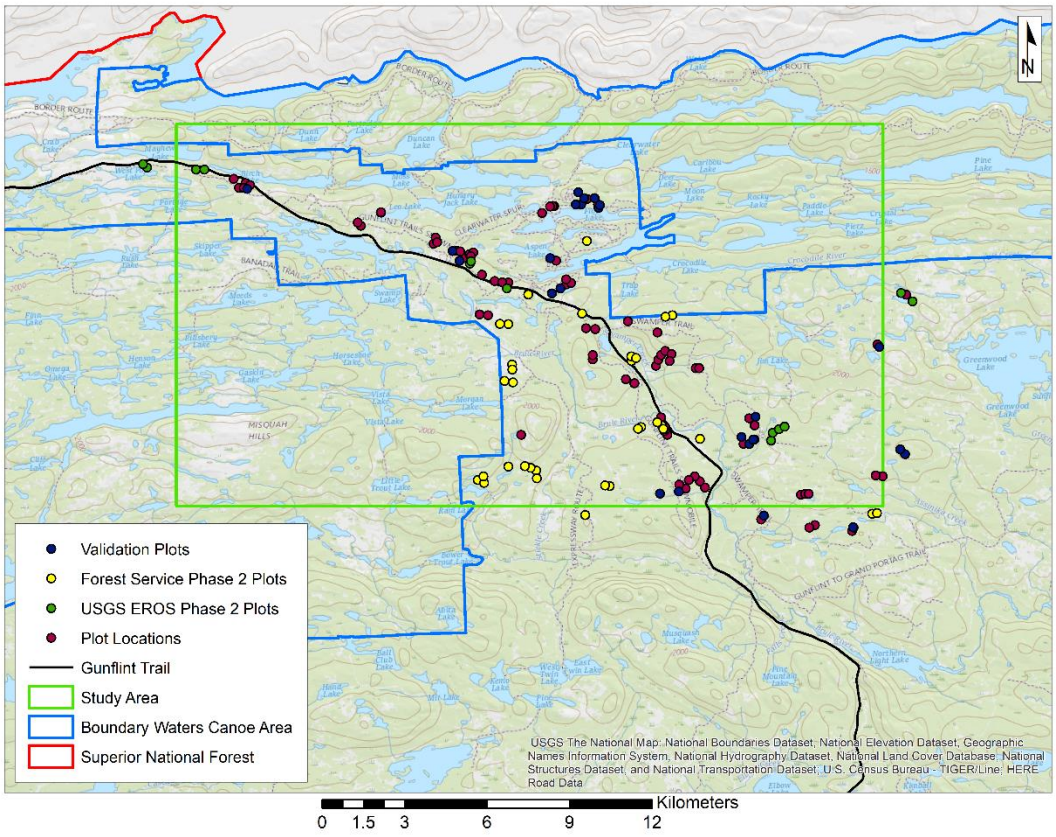

Figure 9. Locations of the validation plots.

species with intermingling canopies separately, the total transect canopy length can be greater than the length of the transect). For each transect that was measured, in each plot, several calculations were made for each individual tree species that was present along the transect (See Table 2). For each species, those calculations were made for several size classes (See Table 3).

Calculations made for each Cover Line Intercept Transect
Number of Intercepts
Total Amount of Canopy Intercepting the Transect
Average Amount of Canopy per Intercept
Percentage of the Total Amount of Canopy Intercepting the Transect
Percent Cover

Table 2. A list of calculations made for each tree species in each transect.

Figure 10. Photo taken along transect 1 of plot 27 by Kurtis Nelson, USGS EROS, July 2016. Note the dense birch and balsam fir understory.

Name	Description
All Trees	All Trees Combined
Very Large Trees	Trees ≥ 80 cm dbh
Large Trees	Trees 50 cm - < 80 cm dbh
Medium Trees	Trees 25 cm - < 50 cm dbh
Pole Trees	Trees 12.5 cm - < 25 cm dbh
Saplings	Trees 2.5 cm - < 12.5 cm dbh
Seedlings	Trees < 2.5 cm dbh
Pole Trees and Larger	Overstory Trees
Saplings and Smaller	Understory Trees

Table 3. Tree size classes (dbh = diameter at breast height).

Figure 11. Photo taken along transect 2 of plot 31 by Birgit Peterson, USGS EROS, May 2017. Note the dead lower limbs on the balsam fir, which add to fuel loads.

Lidar Data and Processing

The Minnesota Department of Natural Resources acquired airborne lidar data over an area that included Superior National Forest in the spring of 2011. The lidar data were originally obtained for the Minnesota Elevation Mapping Project Arrowhead Project Area with funding from the Clean Water Fund of the Clean Water, Land and Legacy Amendment, The Natural Resources Conservation Service, The National Parks Service, and Lake County, MN. The vendor that collected the lidar data was Woolpert, Inc. (Minnesota Department of Natural Resources 2014). The data were made available to the public through the Minnesota Geospatial Information Office website, at

https://gisdata.mn.gov/dataset/elev-lidar-arrowhead2011 (Minnesota Department of Natural Resources 2014).

The lidar data for Cook County, Minnesota (Figure 12) were obtained as laz files (a compressed format) with the points already classified. Figure 13 displays a portion

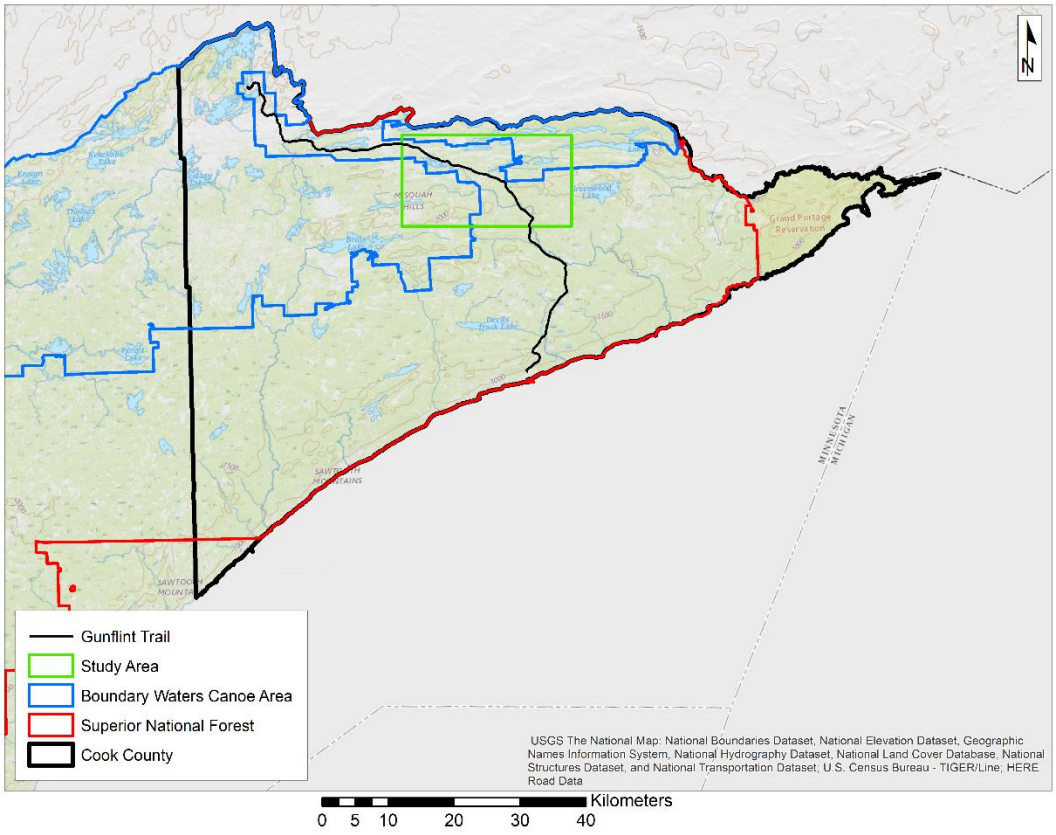

Figure 12. Cook County, Minnesota.

of the original point cloud and Figure 14 shows a profile view from the point cloud. The data are being displayed with the program LP360 (GeoCue Group Inc. 2017). The lidar data were not particularly dense. The Cook County, Minnesota lidar data were collected with a nominal pulse spacing of 1.5 (Minnesota Department of Natural Resources 2014), which equates to 0.44 points per square meter (Renslow 2012).

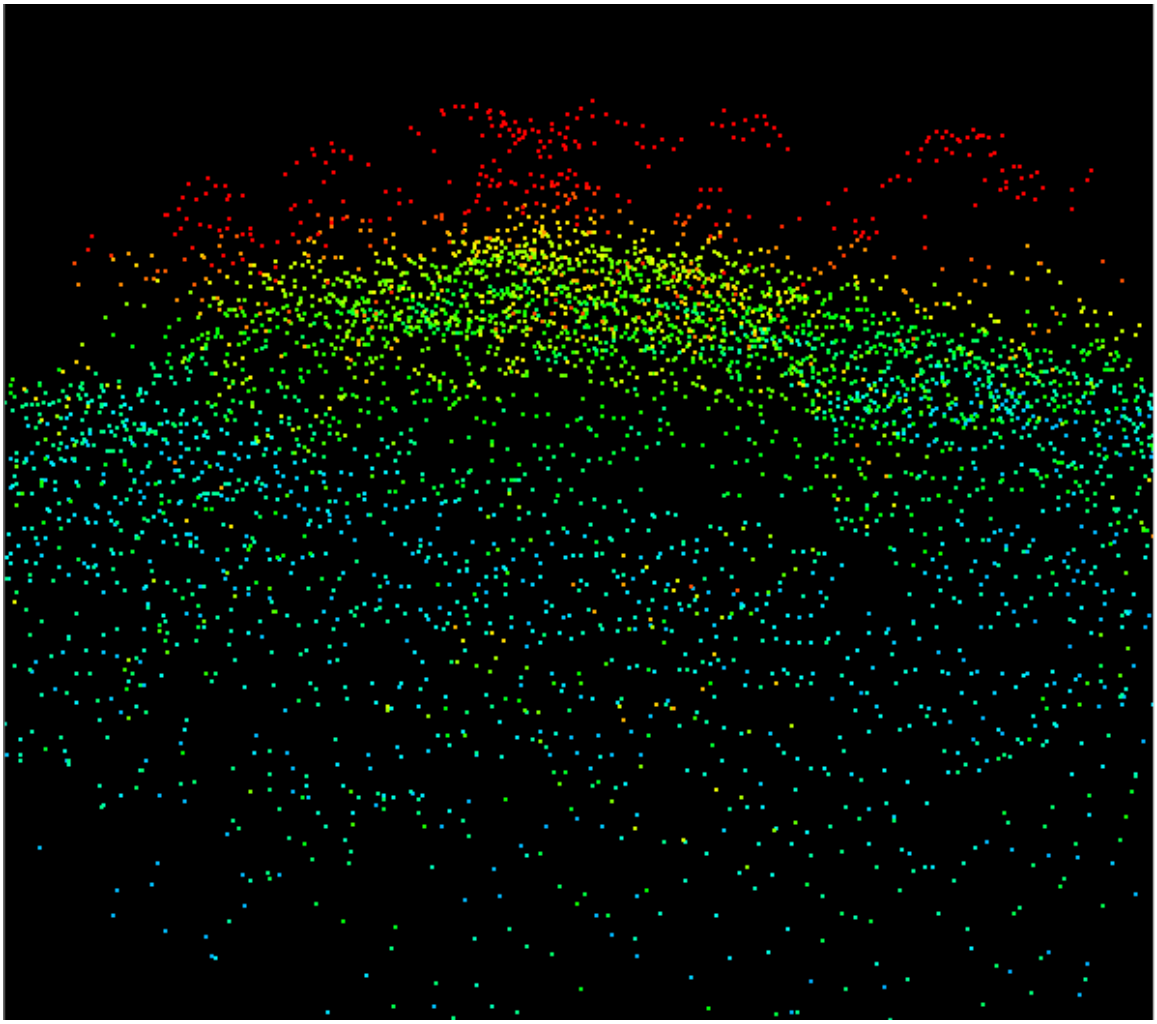

Figure 13. A portion of the original point cloud with the points colored by elevation. Blue points are lowest in elevation and red points are highest.

LAStools (Isenburg 2017) was used to process the lidar data. The lidar data were normalized into a height above ground format using the lasheight function, which sets the elevation of the ground points to 0 and non-ground points to an elevation "that equals their relative height above (or below) the ground," (Isenburg 2017) (Figure 15). From the height normalized point cloud, several lidar metrics were derived (Table 4) using the lascanopy function. The lidar metrics included height percentiles, covers, and relative

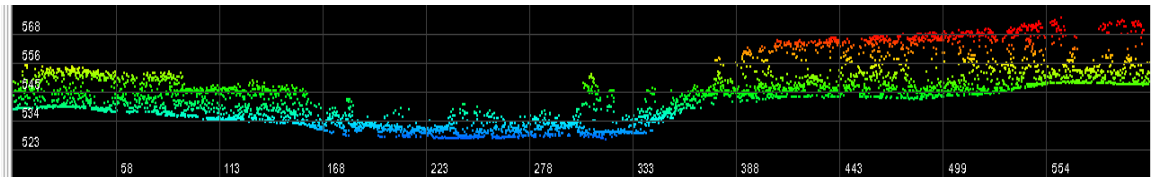

Figure 14. A profile view from the original point cloud. Points are colored by elevation.

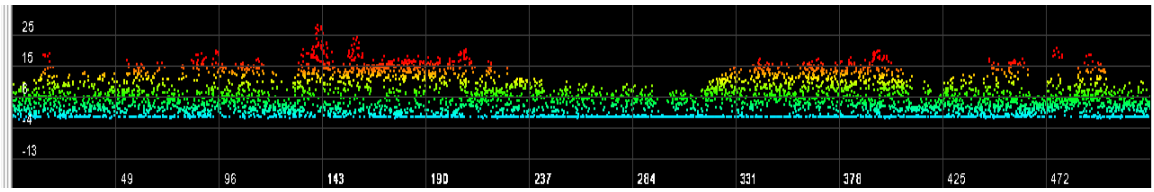

Figure 15. A profile view from the normalized point cloud, notice the flat ground surface. Points are colored by elevation.

densities. LAStools (Isenburg 2017) calculates cover as "the number of first returns above the cover cutoff divided by the number of all first returns and output as a percentage." Relative density is the number of points in a certain height range divided by all points in the same horizontal area (Isenburg 2017).

A height cutoff of 2 meters was used when calculating the covers and height percentiles. Classes 7 and 10, noise and rail, were removed during the lidar metric processing. The lidar metrics were output in the Imagine raster format (.img) at a spatial resolution of 30 meters for all of Cook County, Minnesota. Additionally, the cover, relative density, and height percentile lidar metrics were derived at each plot location with a radius of 20 meters. These plot level metrics would be used as predictor variables in regression analysis (Table 4).

Lidar Metric Name	Divisions
Height Percentile	5^{th}, 10^{th}, 20^{th}, 30^{th}, 40^{th}, 50^{th}, 60^{th}, 70^{th}, 80^{th}, 90^{th}, 95^{th}, 99^{th}
Cover	Total Cover, Cover > 4 m, Cover > 6 m, Cover > 8 m, Cover > 10 m, Cover > 12 m, Cover > 14 m, Cover > 16 m, Cover > 18 m, Cover > 20 m, Cover > 25 m, Cover > 30 m, Cover > 35 m, Cover > 40 m
Relative Density	0 to 2 m, 2 to 4 m, 4 to 6 m, 6 to 8 m, 8 to 10 m, 10 to 12 m, 12 to 14 m, 14 to 16 m, 16 to 18m, 18 to 20 m, 20 to 25 m, 25 to 30 m, 30 to 35 m, 35 to 40 m, 40 to 100 m
Maximum Height	
Average Height	
Standard Deviation	
Average Vegetation Height	

Table 4. Lidar metrics derived from the normalized point cloud. The shaded metrics were used as predictor variables in regression analysis.

Regression Analysis

Regression analysis was used to compare the lidar metrics to the mean amount of understory canopy measured at each plot that had both phase 1 and phase 2 cover line intercept measurements. This amounted to 45 total plots. The mean amount of understory canopy was calculated by summing the total amount of canopy measured for all transects measured at a plot divided by the number of transects measured at that plot during the phase 2 field data collection effort.

According to Brase and Brase (2012), regression analysis works well for data that are continuous. Thus, regression analysis was used in this investigation. Initially analysis was attempted with linear regression. When results were less than satisfactory, multiple regression was attempted.

The goal of linear regression is to fit a straight line to sample data, in order to compare two variables. The two variables are normally plotted onto a cartesian graph with the x variable being the predictor variable (the lidar metric) and the y variable being

the response variable (the measured amount of understory cover). A line can be defined by the following equation:

$$y = a + bx$$

where y is the response variable, a is the intercept on the y axis, b is the slope of the line, and x is the predictor variable. In order obtain the equation for the line that best fits the data, the least squares method is used. The least squares method attempts to make the difference between the measured y value and the predicted y value to be as small as possible over all y values in the data set. The best fit line will pass through a point defined by the mean value of the y variables and the mean value of the x variables. The best fit, or least squares, line can be calculated in two steps. Step one is to compute the slope of the line (b), and step two is to compute the intercept (a). The formula to determine the slope is:

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

where n is the sample size, $\sum xy$ is the sum of all x and y pairs multiplied by each other, $\sum x$ is the sum of all x values, $\sum y$ is the sum of all y values, $\sum x^2$ is the sum of all squared values of x, and $(\sum x)^2$ is the sum of all x values squared. The intercept can be determined by the equation:

$$a = \bar{y} - b\bar{x}$$

where $\bar{y}$ is the mean value of the y variables and $\bar{x}$ is the mean value of the x variables. The least squares line can be used to predict unknown values of the response variable, based upon a value for the predictor variable. The strength of the relationship between two variables can be measured using the sample correlation coefficient (r). The following equation can be used to determine the sample correlation coefficient:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2} - \sqrt{n\sum y^2 - (\sum y)^2}}$$

where $\sum y^2$ is the sum of all squared values of y, and $(\sum y)^2$ is the sum of all y values squared. The sample correlation coefficient ranges between 0 and 1 or 0 and -1, with 0 being no correlation between the variables, 1 being a perfect positive correlation (as x increases, so does y), and -1 being a perfect negative correlation (as x increases, y decreases). Another useful statistic is the coefficient of determination (r^2). The coefficient of determination is a way to measure how much variation is actually explained by the least squares line equation, with the remaining variation being because of random chance, or other variables. The coefficient of determination is obtained by simply squaring the correlation coefficient (Brase and Brase 2012, Crawley 2007).

In multiple regression, more than one predictor variable is used to predict the response variable. The equation for the relationships between the variables is:

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

where y is again the response variable, b_0 is the constant term, b_1, b_2, …, b_k are numerical coefficients, and x_1, x_2, …, x_k are the specific predictor variables. As with linear regression, the least squares criterion for multiple regression seeks to minimize the difference between the measured and predicted y value through the following summation:

$$\sum [y_i - (b_0 + b_1x_{1i} + b_2x_{2i} + \cdots + b_kx_{ki})]^2$$

where $y_i$is the *i*th value of y, x_{1i} is the *i*th value of the x_1 predictor variable, x_{2i} is the *i*th value of the x_2 predictor variable, …, and x_{ki} is *i*th value of the x_k predictor variable. To minimize the summation, the ideal values of the coefficients are determined through a series of normal equations. The number of normal equations that needs to solved is k+ 1,

where k is the number of predictor variables. For example, if there were two predictor variables, the equation to determine the predicted y value would be:

$$y = b_0 + b_1 x_1 + b_2 x_2$$

and three normal equations would need to be solved:

$$\sum y_i = nb_0 + b_1 \left(\sum x_{1i}\right) + b_2 \left(\sum x_{2i}\right)$$

$$\sum x_{1i} y_i = b_0 \left(\sum x_{1i}\right) + b_1 \left(\sum {x_{1i}}^2\right) + b_2 \left(\sum x_{1i} x_{2i}\right)$$

$$\sum x_{2i} y_i = b_0 \left(\sum x_{2i}\right) + b_1 \left(\sum x_{1i} x_{2i}\right) + b_2 \left(\sum {x_{2i}}^2\right)$$

Using actual data values for the x and y variables allows for the determination of the values of b_0, b_1, and b_2. Adding further predictor variables increases the complexity of the calculations. Similar to linear regression, the coefficient of multiple determination (R^2) is used to assess how well the equation fits. The coefficient of multiple determination can be calculated using the following formula:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

where y_i is the measured value for the *i*th measurement of y, $\hat{y}_i$ is the predicted value for the *i*th measurement of y, and $\bar{y}$ is again, the mean value of y (Brase and Brase 2012, Mendenhall and Sincich 2016).

To assist with the statistical calculations, the statistics software R (R Core Team 2016) was used. When the initial linear regression analysis produced unsatisfactory results, multiple regression was used. It was decided that a maximum of five predictor variables would be used to help keep the model meaningful. Originally, a sort of trial and error approach was used, using R's linear model and update functions (R Core Team 2016). Later the leaps package (Lumley 2009) was used to assist with the selection of

significant predictor variables. "leaps() performs an exhaustive search for the best subsets of the variables in x for predicting y in linear regression, using an efficient branch-and-bound algorithm," (Lumley 2009). The leaps package (Lumley 2009) also allows for the specification of the maximum number of desired predictor variables. As with the trial and error effort, a setting of five maximum predictor variables was used.

In the case of both the trial and error effort and the more automated approach, the linear model and update functions were used to reduce the five predictor variables down to the three most significant variables. Care was taken not to choose variables that were highly corelated with each other.

Rasterizing the Understory Cover Model

The resulting equation from regression analysis was used to map the model spatially. ArcMap's (Esri Inc. 1999-2017) raster calculator tool was used to mathematically combine the selected lidar metrics and generate a raster of the estimated amount of understory canopy. The understory canopy model was then converted to a percent understory cover model, using the raster calculator tool.

Results and Discussion

Field Data Collection

The measured amount of understory canopy for each of the 45 plots used for analysis is summarized in Table 5. Plot 121 was excluded from the analysis, as it did not have phase 1 measurements.

Table 5. Plot totals of understory cover measurements. (U. Canopy = the understory canopy measured in meters, plots highlighted in green were measured by the EROS team, plots highlighted in yellow were measured by the U.S. Forest Service, plots highlighted in red were excluded from the analysis, "-" indicates the transect was not measured.)

Plot 13	U. Canopy	% Cover	Plot 14	U. Canopy	% Cover
Trans 1	33.38 m	165.88%	Trans 1	35.30 m	175.43%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	33.38 m	165.88%	Mean	35.30 m	175.43%
Plot 16	U. Canopy	% Cover	Plot 26	U. Canopy	% Cover
Trans 1	32.28 m	160.43%	Trans 1	28.19 m	140.13%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	32.28 m	160.43%	Mean	28.19 m	140.13%
Plot 27	U. Canopy	% Cover	Plot 62	U. Canopy	% Cover
Trans 1	19.26 m	95.74%	Trans 1	40.14 m	199.52%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	19.26 m	95.74%	Mean	40.14 m	199.52%
Plot 79	U. Canopy	% Cover	Plot 80	U. Canopy	% Cover
Trans 1	9.60 m	47.72%	Trans 1	15.24 m	75.75%
Trans 2	7.77 m	38.63%	Trans 2	13.11 m	65.14%
Trans 3	15.70 m	87.11%	Trans 3	12.19 m	60.60%
Trans 4	17.53 m	87.11%	Trans 4	11.28 m	56.05%
Mean	12.65 m	62.87%	Mean	12.95 m	64.38%
Plot 84	U. Canopy	% Cover	Plot 89	U. Canopy	% Cover
Trans 1	15.39 m	76.50%	Trans 1	10.06 m	49.99%
Trans 2	8.08 m	40.15%	Trans 2	-	-
Trans 3	7.47 m	37.12%	Trans 3	-	-
Trans 4	21.18 m	105.29%	Trans 4	-	-
Mean	13.03 m	64.76%	Mean	10.06 m	49.99%
Plot 90	U. Canopy	% Cover	Plot 92	U. Canopy	% Cover
Trans 1	3.20 m	15.91%	Trans 1	26.18 m	130.13%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	3.20 m	15.91%	Mean	26.18 m	130.13%

Table 5. Continued.

Plot 95	U. Canopy	% Cover	Plot 98	U. Canopy	% Cover
Trans 1	7.01 m	34.84%	Trans 1	1.37 m	6.82%
Trans 2	8.84 m	43.93%	Trans 2	8.69 m	43.17%
Trans 3	3.96 m	19.69%	Trans 3	10.82 m	53.78%
Trans 4	6.71 m	33.33%	Trans 4	5.94 m	29.54%
Mean	6.63 m	32.95%	Mean	6.71 m	33.33%
Plot 99	U. Canopy	% Cover	Plot 103	U. Canopy	% Cover
Trans 1	1.68 m	8.33%	Trans 1	9.45 m	46.96%
Trans 2	2.90 m	14.39%	Trans 2	19.96 m	99.23%
Trans 3	4.88 m	24.24%	Trans 3	20.88 m	103.77%
Trans 4	2.13 m	10.60%	Trans 4	28.80 m	143.16%
Mean	2.90 m	14.39%	Mean	19.77 m	98.28%
Plot 104	U. Canopy	% Cover	Plot 105	U. Canopy	% Cover
Trans 1	6.10 m	30.30%	Trans 1	8.53 m	42.42%
Trans 2	9.14 m	45.45%	Trans 2	11.58 m	57.57%
Trans 3	11.89 m	59.08%	Trans 3	23.32 m	115.89%
Trans 4	8.53 m	42.42%	Trans 4	3.81 m	18.94%
Mean	8.92 m	44.31%	Mean	11.81 m	58.70%
Plot 106	U. Canopy	% Cover	Plot 107	U. Canopy	% Cover
Trans 1	6.86 m	34.09%	Trans 1	13.01 m	64.69%
Trans 2	10.06 m	49.99%	Trans 2	8.02 m	39.84%
Trans 3	7.62 m	37.87%	Trans 3	3.05 m	15.15%
Trans 4	3.51 m	17.42%	Trans 4	14.17 m	70.44%
Mean	7.01 m	34.84%	Mean	9.56 m	47.53%
Plot 108	U. Canopy	% Cover	Plot 121	U. Canopy	% Cover
Trans 1	4.11 m	20.45%	Trans 1	20.97 m	104.22%
Trans 2	2.90 m	14.39%	Trans 2	-	-
Trans 3	12.65 m	62.87%	Trans 3	-	-
Trans 4	9.30 m	46.20%	Trans 4	-	-
Mean	7.24 m	35.98%	Mean	20.97 m	104.22%
Plot 122	U. Canopy	% Cover	Plot 141	U. Canopy	% Cover
Trans 1	19.23 m	95.59%	Trans 1	3.51 m	17.42%
Trans 2	-	-	Trans 2	11.28 m	56.05%
Trans 3	-	-	Trans 3	10.06 m	49.99%
Trans 4	-	-	Trans 4	3.51 m	17.42%
Mean	19.23 m	95.59%	Mean	7.09 m	35.22%
Plot 142	U. Canopy	% Cover	Plot 143	U. Canopy	% Cover
Trans 1	15.24 m	75.77%	Trans 1	5.33 m	26.49%
Trans 2	18.14 m	90.14%	Trans 2	8.68 m	43.16%
Trans 3	21.34 m	106.06%	Trans 3	8.99 m	44.69%
Trans 4	20.72 m	102.98%	Trans 4	14.02 m	69.70%
Mean	18.86 m	93.74%	Mean	9.26 m	46.01%

Table 5. Continued.

Plot 150	U. Canopy	% Cover	Plot 151	U. Canopy	% Cover
Trans 1	2.50 m	12.42%	Trans 1	6.31 m	31.36%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	2.50 m	12.42%	Mean	6.31 m	31.36%
Plot 152	U. Canopy	% Cover	Plot 153	U. Canopy	% Cover
Trans 1	18.29 m	90.89%	Trans 1	17.34 m	86.20%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	18.29 m	90.89%	Mean	17.34 m	86.20%
Plot 168	U. Canopy	% Cover	Plot 210	U. Canopy	% Cover
Trans 1	7.31 m	36.36%	Trans 1	15.39 m	76.50%
Trans 2	12.19 m	60.60%	Trans 2	4.11 m	20.45%
Trans 3	15.24 m	75.75%	Trans 3	8.99 m	44.69%
Trans 4	9.75 m	48.48%	Trans 4	2.59 m	12.88%
Mean	11.13 m	55.30%	Mean	7.77 m	38.63%
Plot 211	U. Canopy	% Cover	Plot 212	U. Canopy	% Cover
Trans 1	5.79 m	28.78%	Trans 1	13.87 m	68.93%
Trans 2	1.37 m	6.82%	Trans 2	9.75 m	48.48%
Trans 3	2.44 m	12.12%	Trans 3	16.76 m	83.32%
Trans 4	0.61 m	3.03%	Trans 4	10.52 m	52.26%
Mean	2.55 m	12.69%	Mean	12.73 m	63.25%
Plot 214	U. Canopy	% Cover	Plot 222	U. Canopy	% Cover
Trans 1	15.24 m	75.75%	Trans 1	17.22 m	85.59%
Trans 2	17.53 m	87.11%	Trans 2	-	-
Trans 3	15.24 m	75.75%	Trans 3	-	-
Trans 4	18.75 m	93.17%	Trans 4	-	-
Mean	16.69 m	82.94%	Mean	17.22 m	85.59%
Plot 223	U. Canopy	% Cover	Plot 228	U. Canopy	% Cover
Trans 1	18.90 m	93.92%	Trans 1	10.97 m	54.54%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	18.90 m	93.92%	Mean	10.97 m	54.54%
Plot 229	U. Canopy	% Cover	Plot 230	U. Canopy	% Cover
Trans 1	5.49 m	27.27%	Trans 1	6.10 m	30.30%
Trans 2	-	-	Trans 2	-	-
Trans 3	-	-	Trans 3	-	-
Trans 4	-	-	Trans 4	-	-
Mean	5.49 m	27.27%	Mean	6.10 m	30.30%

Table 5. Continued.

Plot 239	U. Canopy	% Cover	Plot 240	U. Canopy	% Cover
Trans 1	18.90 m	93.92%	Trans 1	10.67 m	53.02%
Trans 2	19.35 m	96.20%	Trans 2	-	-
Trans 3	13.11 m	65.14%	Trans 3	-	-
Trans 4	18.44 m	91.65%	Trans 4	-	-
Mean	17.45 m	86.73%	Mean	10.67 m	53.02%
Plot 241	U. Canopy	% Cover	Plot 242	U. Canopy	% Cover
Trans 1	15.54 m	77.26%	Trans 1	6.40 m	31.81%
Trans 2	7.62 m	37.87%	Trans 2	-	-
Trans 3	18.90 m	93.92%	Trans 3	-	-
Trans 4	22.71 m	112.86%	Trans 4	-	-
Mean	16.19 m	80.48%	Mean	6.40 m	31.81%
Plot 243	U. Canopy	% Cover	Plot 247	U. Canopy	% Cover
Trans 1	5.18 m	25.75%	Trans 1	12.50 m	62.11%
Trans 2	15.70 m	78.02%	Trans 2	-	-
Trans 3	10.06 m	49.99%	Trans 3	-	-
Trans 4	7.16 m	35.60%	Trans 4	-	-
Mean	9.52 m	47.34%	Mean	12.50 m	62.11%

There was a total of 114 transects measured in the 45 plots. The total amount of understory canopy for those 114 transects was 1,373.55 meters. The amount of understory canopy ranged from 0.61 – 40.14 meters, with a mean of 12.05 meters of understory canopy per transect. This equated to a range of understory cover from 3.03 – 199.52%, with a mean of 59.88% per transect.

For the 45 plots, the total of the mean understory canopy measured was 628.32 meters, with a plot mean of 13.92 meters. The range of mean plot understory canopy was 2.50 – 40.14 meters. In terms of percent understory cover, the mean was 69.40%, with a minimum of 12.42% and a maximum of 199.52%. Both the mean plot understory cover and the minimum plot understory cover were greater than the mean transect understory cover and the minimum transect understory cover. The maximum transect understory cover was equal to the maximum plot understory cover, this was because transect 1 in plot 62 had the greatest amount of understory canopy and was the only transect measured in plot 62.

A consequence of using the cover line intercept method was that the total canopy measured could actually be greater than the length of the transect. This could occur in transects where differing tree species had intermingling canopies, which resulted in a separate measurement for each species.

Lidar Data and Processing

An issue arose during lidar metric production. Striping appeared on the raster outputs of the lidar metrics. It was determined that class 17, which should correspond to bridge deck, was misclassified. In reality, the class 17 points were from overlapping flight lines. Lidar metrics were reprocessed with class 17 dropped in addition to classes 7 and 10. Figure 16 displays the raster output for the 99th height percentile lidar metric.

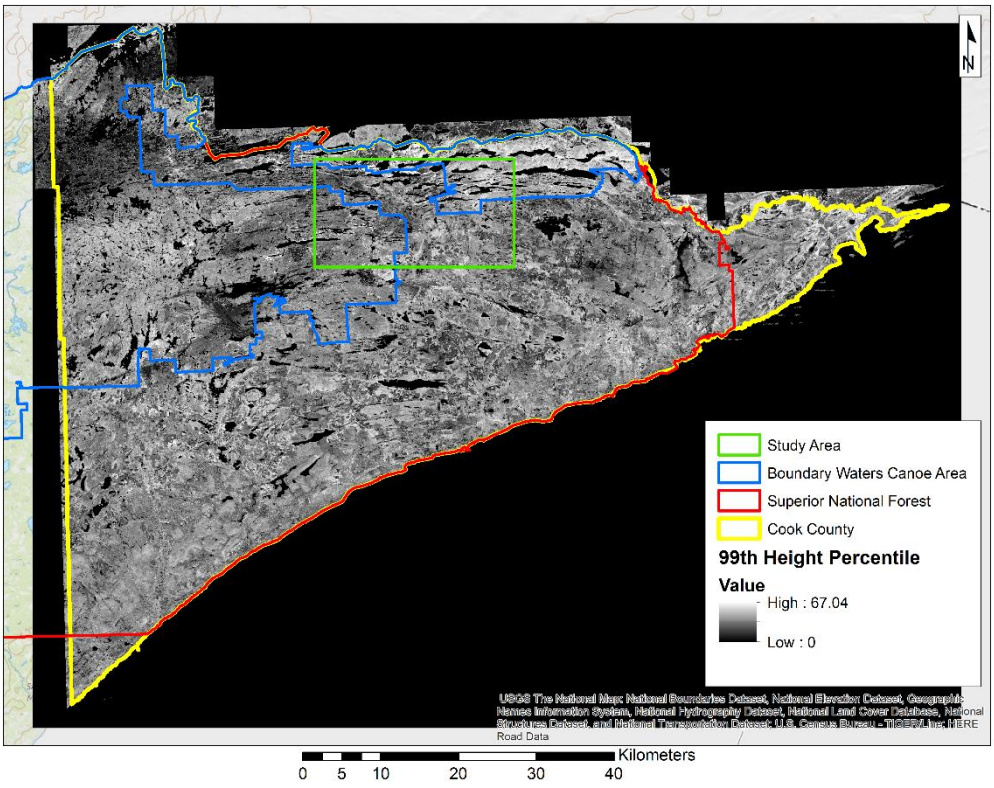

Figure 16. Raster image for the 99th height percentile lidar metric. Output is in meters.

Regression Analysis

In the initial attempt to establish a relationship between the cover line intercept data and the lidar metrics, linear regression was used. The leaps package (Lumley 2009) can output a unique graph, which can be used to select the most highly correlated variables (Figure 17).

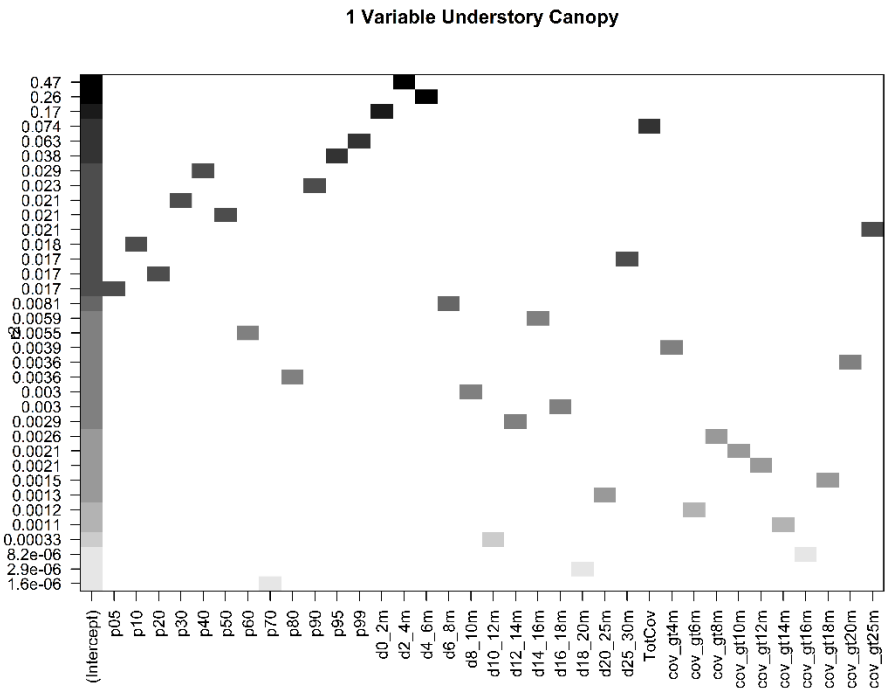

Figure 17. A graph showing the correlation of each lidar metric with the measured understory canopy. The y-axis displays the r^2 values and the x-axis displays the lidar metrics.

The relative density between 2 and 4 meters was the most highly correlated variable with the measured understory canopy (Figure 17). The coefficient of determination (r^2) was 0.4678 (Figures 18 and 19), which equates to a correlation coefficient (r) of 0.6840. It is interesting to note that Kramer et al. (2014) found that the cover between 2 and 4 meters was the best lidar metric for predicting whether, or not,

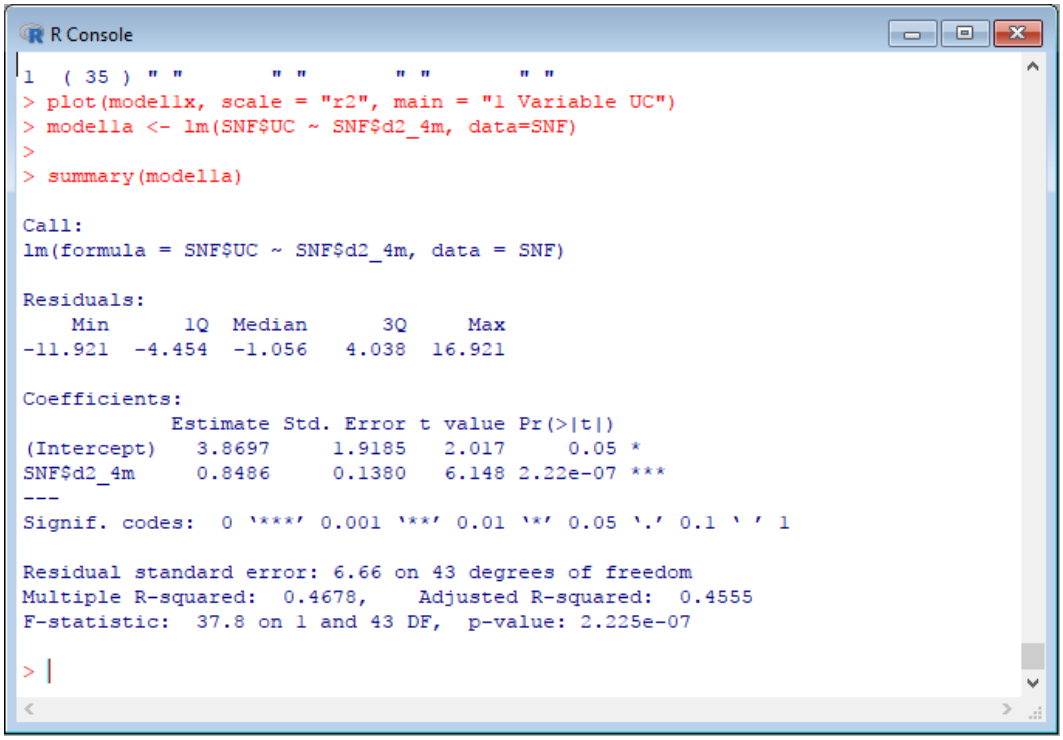

Figure 18. Output from R (R Core Team 2016), showing the linear model results comparing the measured understory canopy with the relative density between 2 and 4 meters lidar metric.

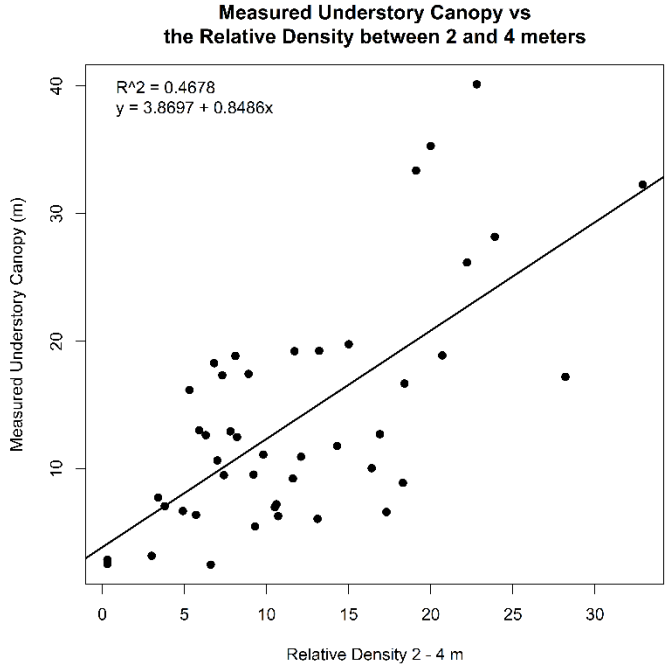

Figure 19. Measured understory canopy vs relative density between 2 and 4 meters.

forest stands in the Plumas National Forest had been treated for ladder fuels. Kramer et al. (2014) did not calculate relative density metrics in their investigation and cover was not calculated in height bins in this investigation.

Multiple regression was used to improve the results. As mentioned above, it was decided that five predictor variables would be the maximum number that would be used, in an effort to keep the model meaningful. The initial trial and error method returned a model with a coefficient of determination of 0.6269, using the following predictor variables: relative density from 2 to 4 meters, relative density from 8 to 10 meters, cover greater than 4 meters, and height of the 50th percentile of returns (Figure 20). It is likely

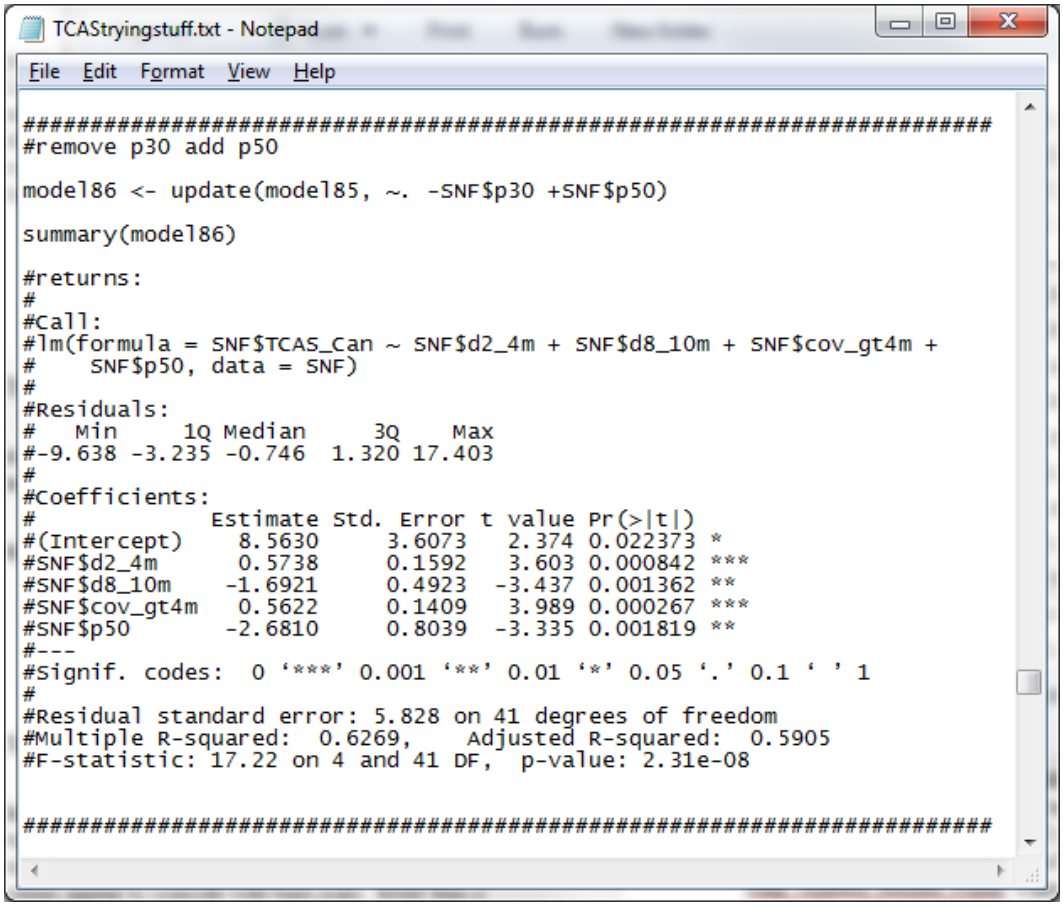

Figure 20. R code and output for the trial and error understory canopy model.

the relative density from 2 to 4 meters metric was having a large impact on the coefficient of determination, since individually that metric had a coefficient of determination approaching 0.47.

For the sake of comparison and improvement, the leaps package (Lumley 2009) was used to generate a model. The parameters were set for a maximum of five predictor variables. The five predictor variables that the leaps package (Lumley 2009) returned were the height of the 50th percentile of returns, relative density between the heights of 4 and 6 meters, relative density between the heights of 8 and 10 meters, total cover, and cover greater than 12 meters (Figure 21). Those five predictor variables were then plugged into the linear model function in R (R Core Team 2016) and the summary function was used to inspect the model (Figure 22).

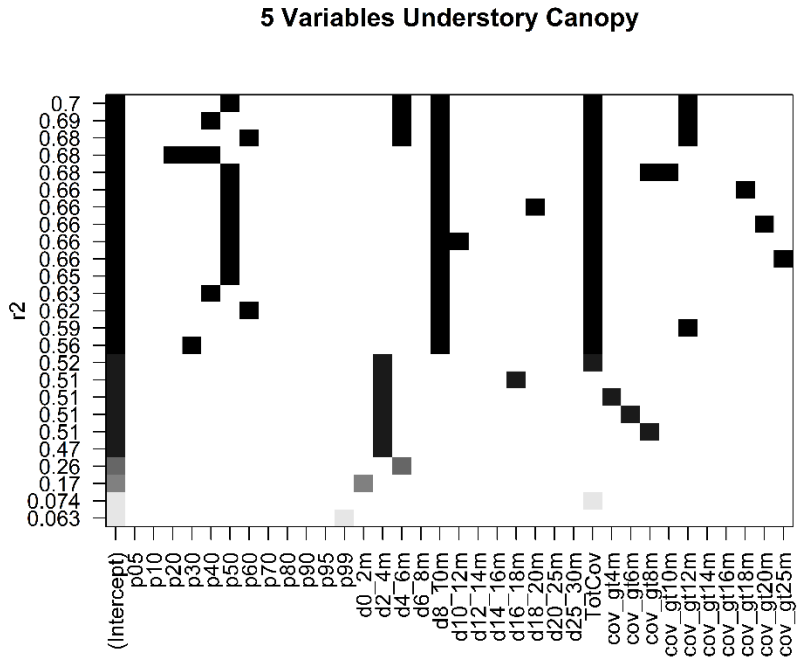

Figure 21. Leaps package (Lumley 2009) graph used for predictor variable selection.

Since the intercept was not significant (Figure 22), the weakest variable, relative density between the heights of 4 and 6 meters, was removed using the update function (Figure 23). Removing the relative density between the heights of 4 and 6 meters metric

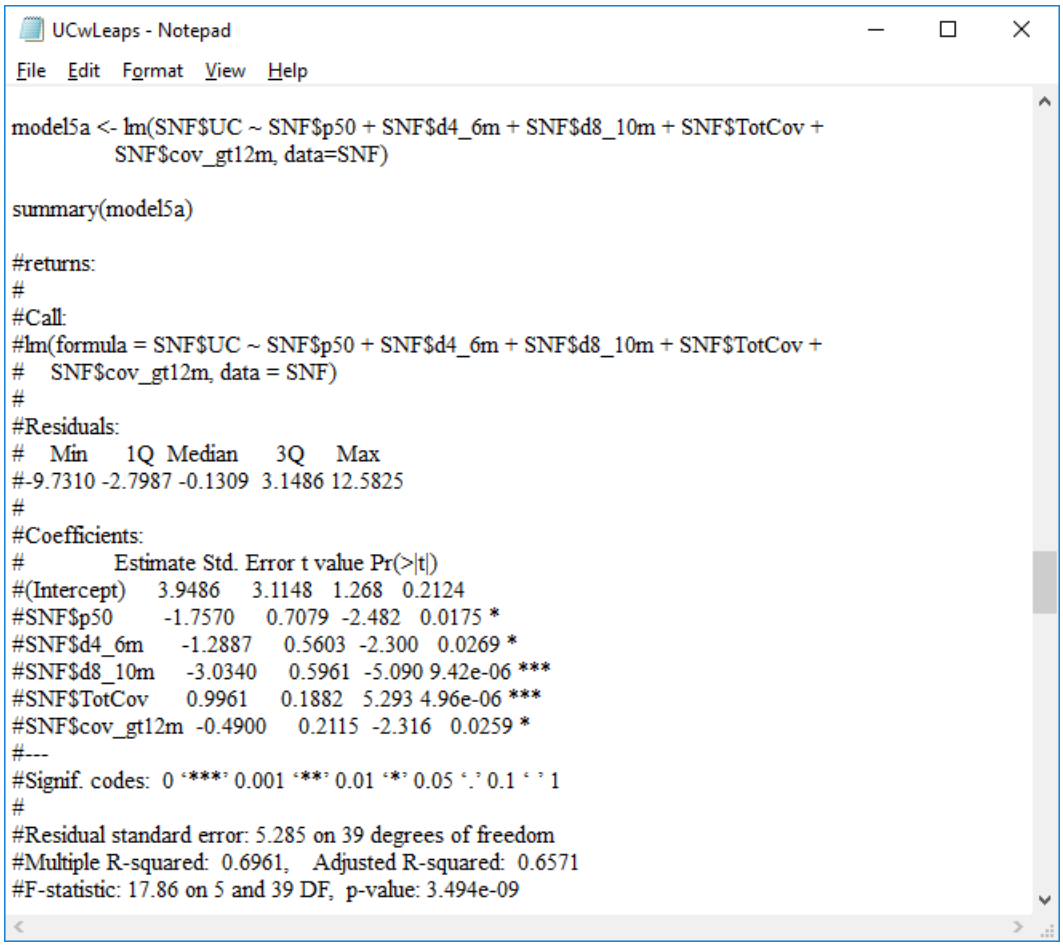

Figure 22. Output for the five identified predictor variables.

caused the cover greater than 12 meters metric to no longer be significant, so it was removed (Figure 24). The resulting model had three predictor variables, the height of the 50th percentile of returns, relative density between the heights of 8 and 10 meters, and total cover. These three variables were significant at a P-value of 0.001. The intercept that was significant at a P-value of 0.01 (Figure 24). The coefficient of multiple determination for the model was 0.6504, an improvement on the trial and error approach.

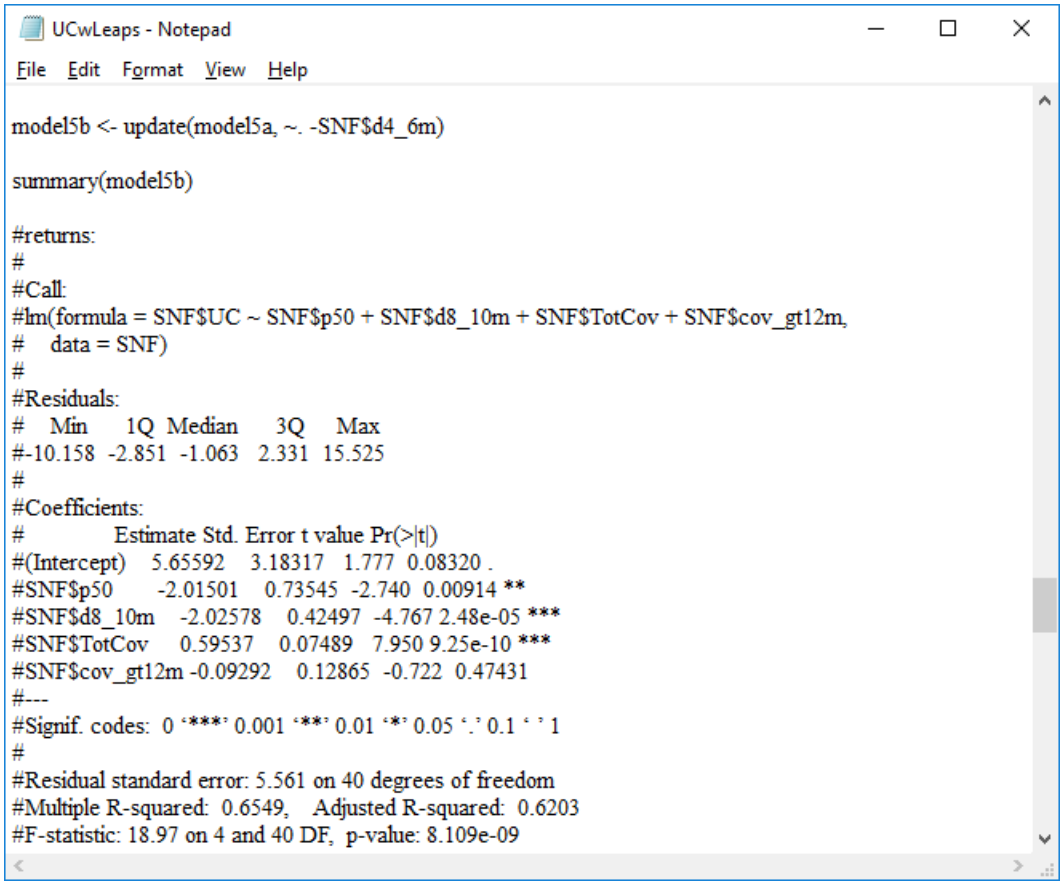

Figure 23. Results from removing the relative density between 4 and 6 meters metric.

Additionally, this model had only three predictor variables, where the trial and error model used four.

The process was repeated using a setting of four maximum predictor variables with the leaps package (Lumley 2009). The final results were the same. Leaps (Lumley 2009) was also run with a three-variable maximum. The output graph indicated that the same final three metrics, the height of the 50th percentile of returns, relative density between the heights of 8 and 10 meters, and total cover would produce the best model (Figure 25).

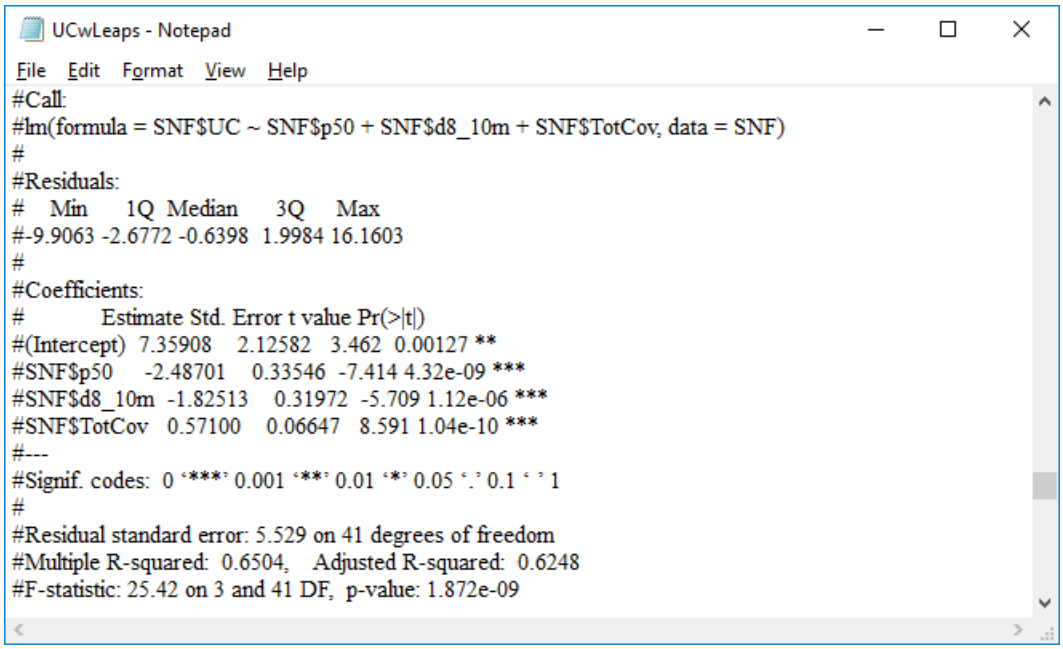

Figure 24. Output after removing the cover greater than 12 meters metric.

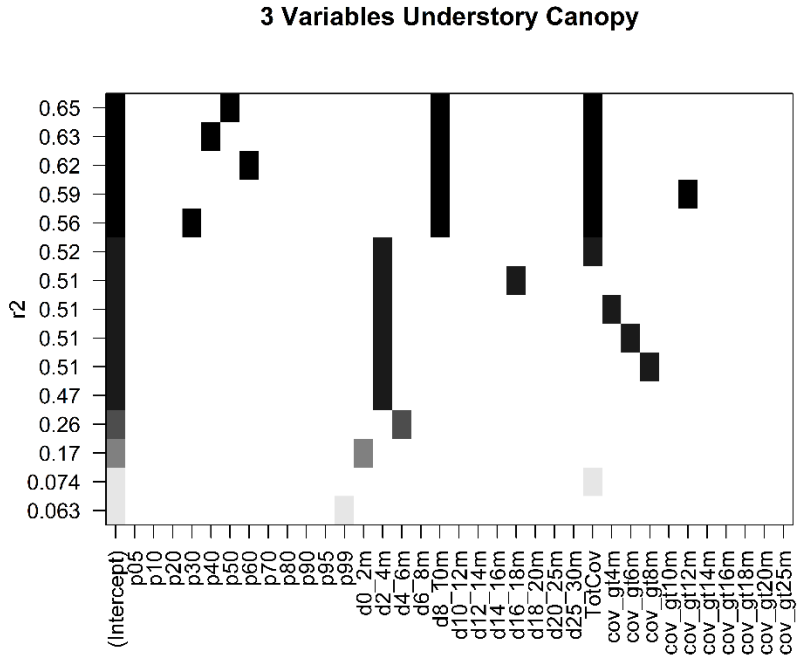

Figure 25. Leaps graph for three variable selection, confirming the final results of the five-variable effort.

Rasterizing the Understory Cover Model and Initial Inspection

The results from the multiple regression analysis produced an equation for the predicted amount of understory canopy. The equation was:

$$UC = 7.35908 + -2.48701(p50) + -1.82513(d8\_10m) + 0.57100(TotCov)$$

where UC is the understory canopy length in meters, p50 is the height at which 50% of the lidar points are returned (meters), d8_10m is the relative point density between 8 and 10 m (percentage), and TotCov is the total cover metric (percentage). Figures 26, 27, and 28 show the raster images for those lidar metrics.

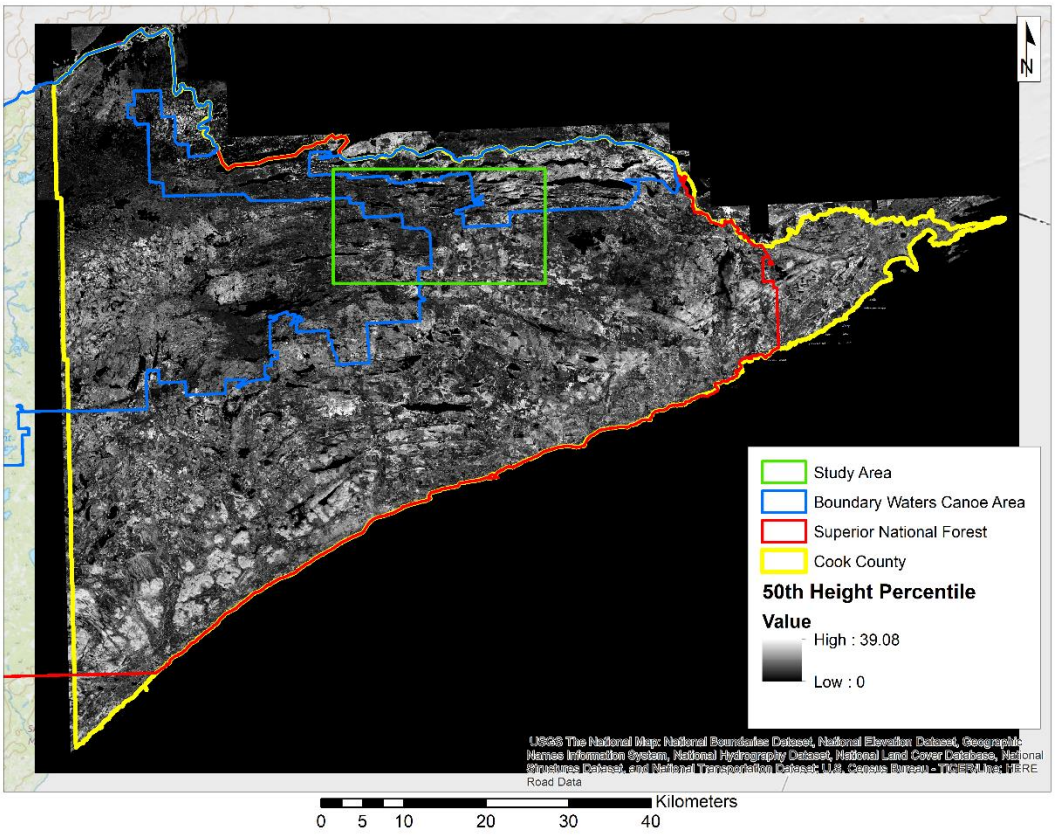

Figure 26. Raster for the height at which 50% of the lidar points are returned metric at 30-meter resolution.

The model's equation and the three raster images for the selected lidar metrics were used to produce a spatial output of the model. ArcMap's (Esri Inc. 1999-2017) raster calculator tool was used to produce the rasterized format of the model (Figure 29). The understory canopy model was then converted into a spatial understory cover model (Figure 30), again using the raster calculator tool.

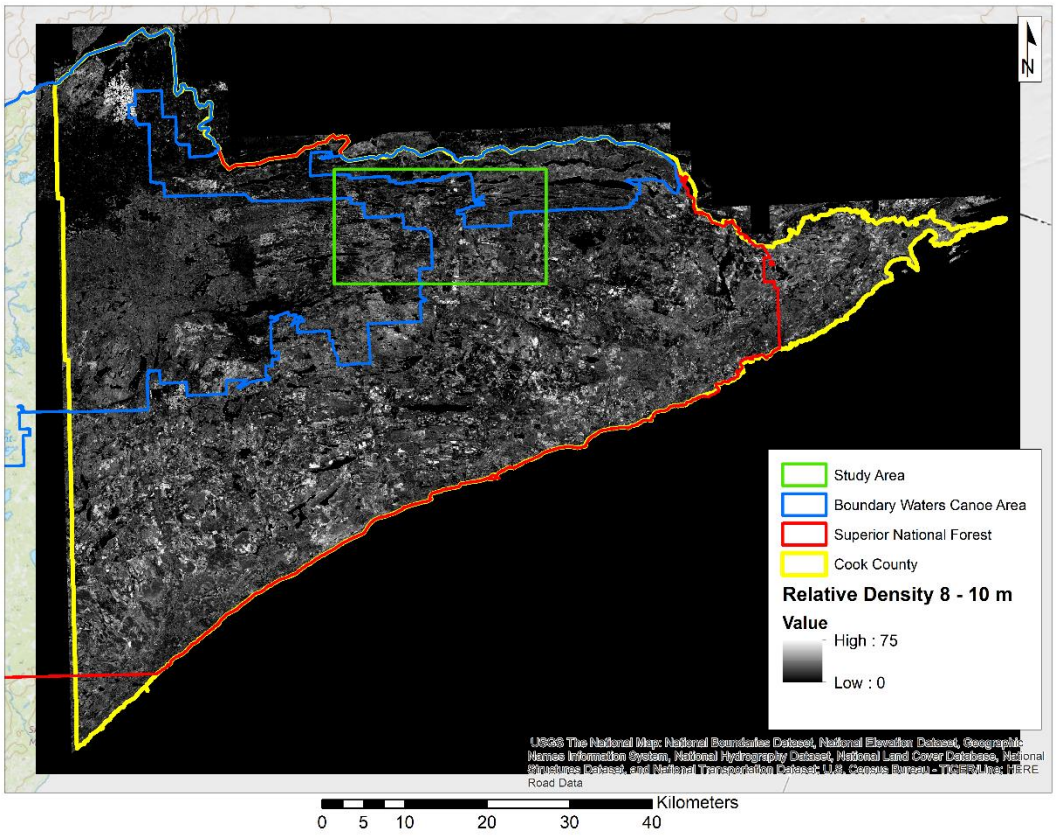

Figure 27. Raster for the relative density between the heights of 8 and 10 meters metric at 30-meter resolution.

The understory cover raster had some interesting values that were not expected. The understory cover model's values ranged from -1,702.02 to 59.4602 meters (Figure 30). However, the high values were not necessarily a surprise, because of the possibility

of having a transect measurement of understory canopy that was longer than the transect using the cover line intercept method.

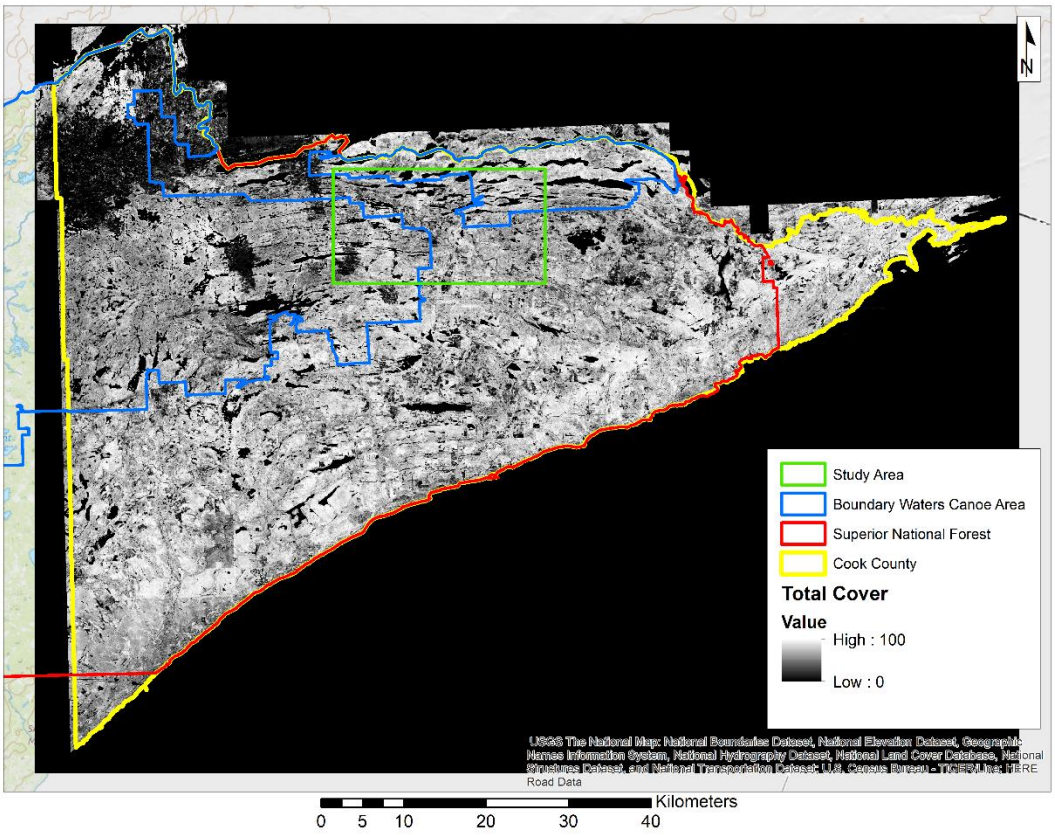

Figure 28. Raster for the total cover lidar metric at 30-meter resolution.

The identify tool was used to inspect several pixels in the understory canopy raster (Figure 29). None of the pixels returned extreme negative values for the amount of understory canopy. Some pixels did return a value that was around -20 meters, however, most were not less than -10 meters. Forty-four random pixels were selected to check the calculations for the model. Each pixel was correct based upon the mathematics.

ArcMap's (Esri Inc. 1999-2017) sample tool was used to generate a table that contained the values of the individual pixels in the understory cover model. While it

should be noted that the total pixel count included all of the rectangular raster, only 287,173 of the 12,042,525 pixels (2.4%) came out with a value of less than 0%

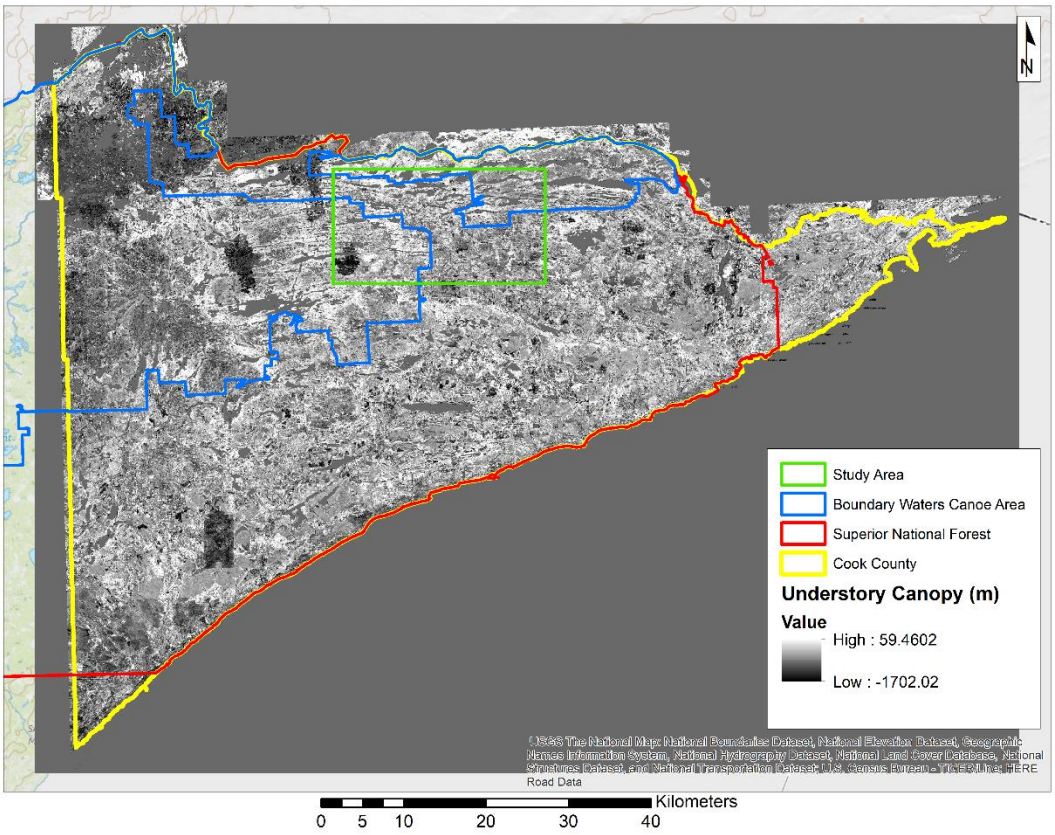

Figure 29. Spatial output of the understory canopy model at 30-meter resolution.

understory cover and only 53,428 pixels (0.4%) had a value of less than -50% understory cover. Only 227 pixels (0.002%) were less than -200% understory cover, which indicated that the extreme negative were rare. On the other end of the spectrum, 1,494,385 pixels (12.4%) were greater than 100% understory cover, with only 57,889 (0.5%) being greater than 200% understory cover, and 1,619 pixels (0.01%) were greater than 250% understory cover. Again, the high values were not as surprising as the low

values, given the cover line intercept methodology. Figure 31 show the histogram and statistics provided from the table.

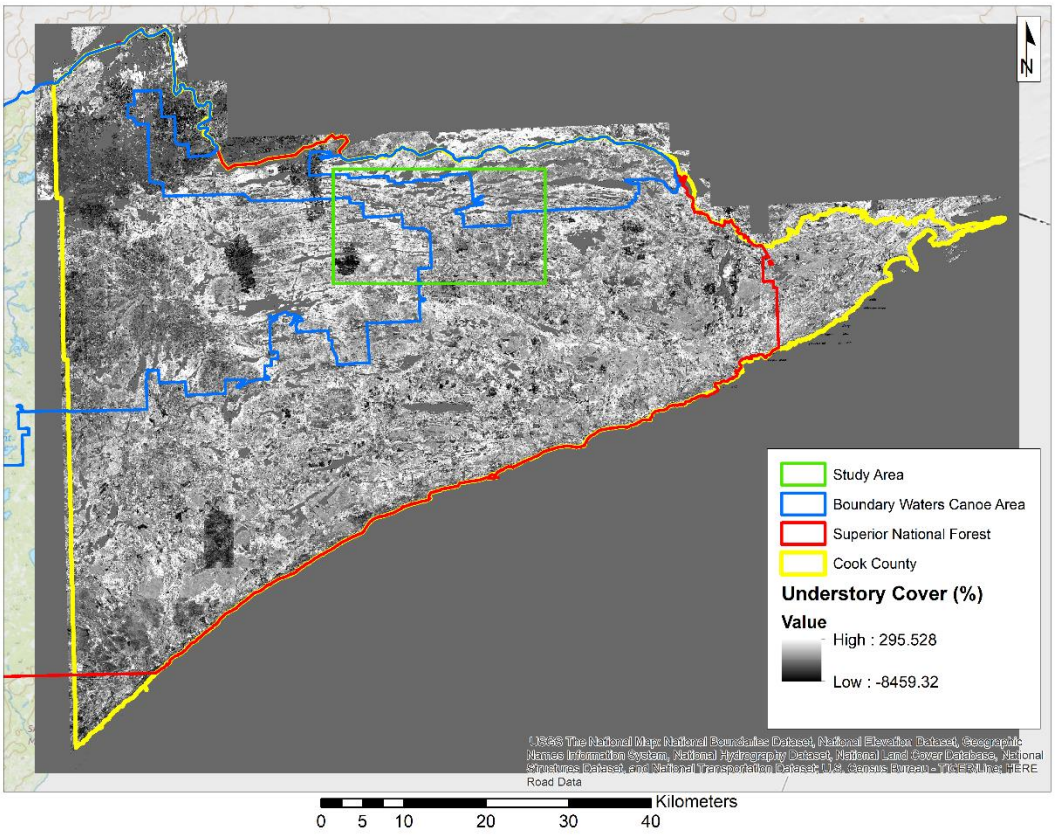

Figure 30. The model expressed as understory cover at 30-meter resolution.

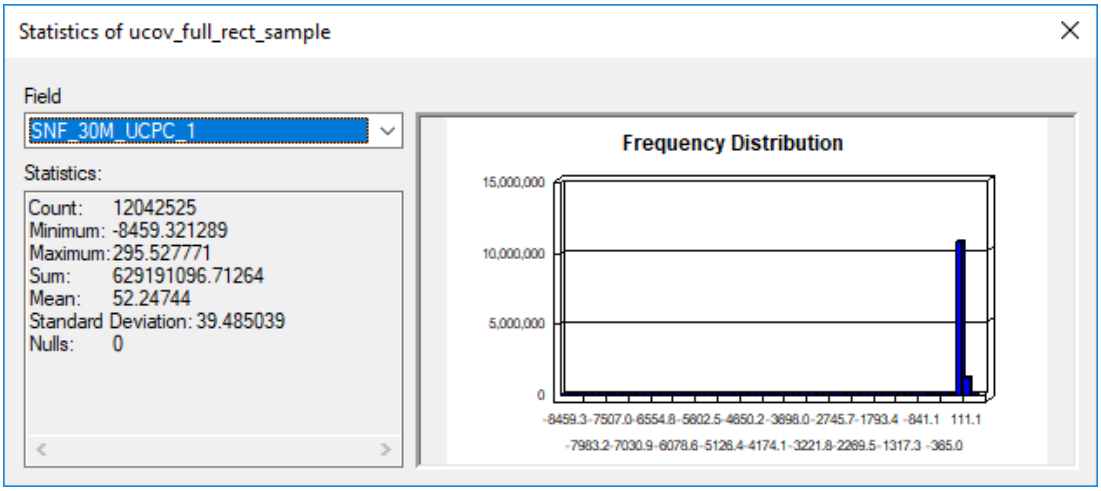

Figure 31. Statistics and histogram of the pixel values for the understory cover model.

Another interesting phenomenon that emerged in the modeled raster was the large areas with low amounts of understory cover (Figure 30). The areas have been compared to Monitoring Trends in Burn Severity data for the years 2002 – 2012 (Monitoring Trends in Burn Severity 2016a, Monitoring Trends in Burn Severity 2016b, Monitoring Trends in Burn Severity 2016c, Monitoring Trends in Burn Severity 2016d, Monitoring Trends in Burn Severity 2016e, Monitoring Trends in Burn Severity 2016f, Monitoring Trends in Burn Severity 2016g, Monitoring Trends in Burn Severity 2016h, Monitoring Trends in Burn Severity 2016i, Monitoring Trends in Burn Severity 2016j, Monitoring Trends in Burn Severity 2016k). Monitoring Trends in Burn Severity is a joint project between the U.S. Geological Survey Earth Resources Observation and Science Center and the USDA Forest Service Geospatial Technology and Applications Center. Monitoring Trends in Burn Severity has mapped the burn severity of all fires in the United States that were larger than 500 acres (200 hectares) in the east and 1,000 acres (400 hectares) in the west since 1984 (Monitoring Trends in Burn Severity 2017). The fire severity maps for the 2006 Cavity Lake, Famine Lake, and Redeye fires and 2007 Ham Lake fire suggest that the areas of low understory cover are areas that had burned relatively recently, when compared to the airborne lidar data collection date (Figure 32).

Further investigation into the smaller areas with low understory cover was necessary, because of the large disturbance area required for Monitoring Trends in Burn Severity mapping. LANDFIRE disturbance data were obtained for the years 1999 – 2012 (LANDFIRE 2012a, LANDFIRE 2012b, LANDFIRE 2012c, LANDFIRE 2012d, LANDFIRE 2012e, LANDFIRE 2012f, LANDFIRE 2012g, LANDFIRE 2012h,

LANDFIRE 2012i, LANDFIRE 2012j, LANDFIRE 2012k, LANDFIRE 2012l, LANDFIRE 2012m, LANDFIRE 2012n). LANDFIRE is a joint program between the

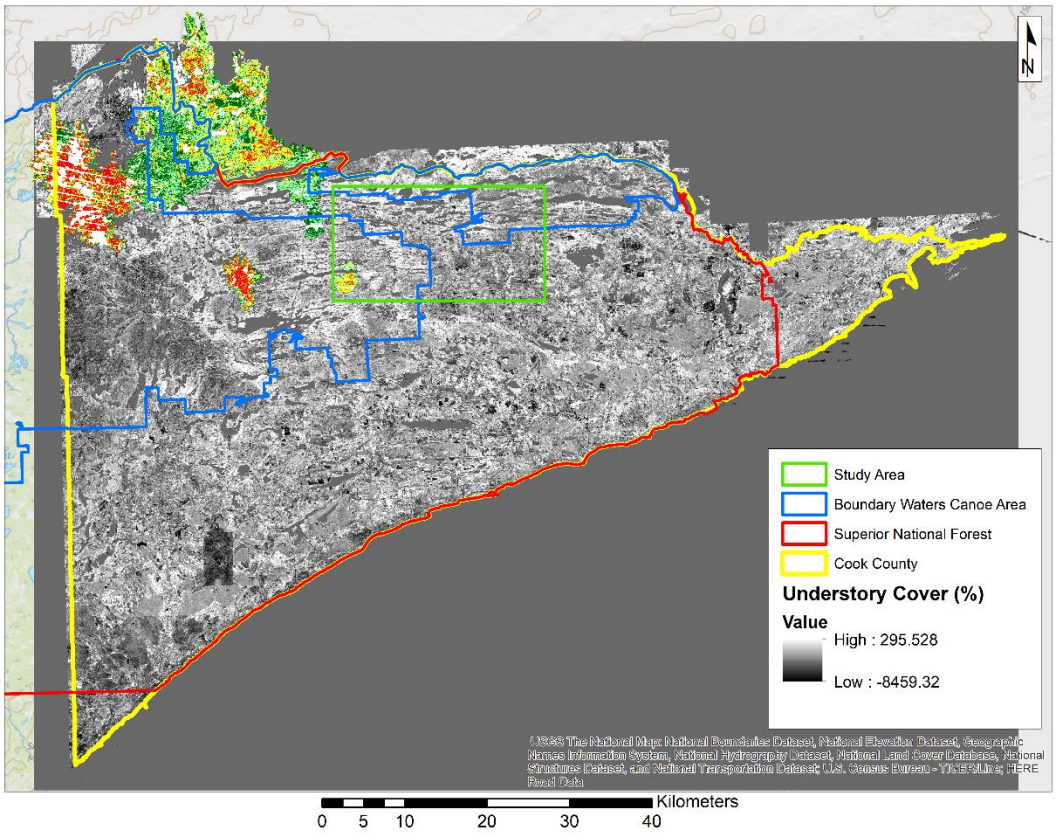

Figure 32. The 2006 Cavity Lake, Famine Lake, and Redeye fires and 2007 Ham Lake fire burn severity maps (Monitoring Trends in Burn Severity 2016e, Monitoring Trends in Burn Severity 2016f) displayed with the understory cover model. Green areas represent low burn severity and red areas represent high burn severity.

U.S. Department of Interior and U.S. Department of Agriculture Forest Service wildland fire management programs that seeks to "provide agency leaders and managers with a common 'all-lands' data set of vegetation and wildland fire/fuels information for strategic fire and resource management planning and analysis," (LANDFIRE 2013). The data were separated out by disturbance type and visually compared to the understory cover model (Figure 33).

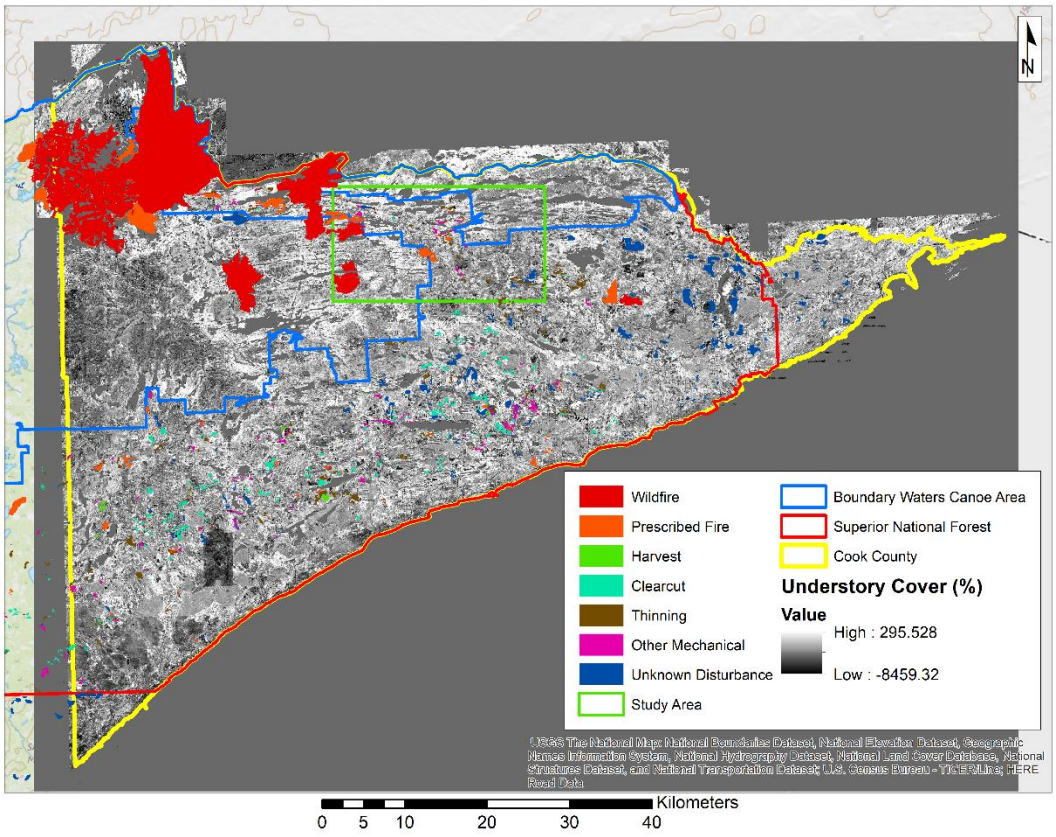

Figure 33. LANDFIRE disturbance data appears to match with many of the areas that are low in understory cover.

There was an area in the southwest portion of the model that did not appear to match any disturbance data, from either Monitoring Trends in Burn Severity or LANDFIRE (Figures 29, 30, 32, and 33). The area was nearly rectangular in shape. Further investigation revealed that it was an area with many overlapping flight lines from the airborne lidar collection. The area was slightly visible in the total cover raster (Figure 28), but not in either of the other two lidar metrics that went into model building (Figures 26 and 27).

Cleaning Up and Rescaling the Understory Cover Model

The spatial understory cover model was trimmed down to the extent of Cook County, Minnesota using the extract by mask tool in ArcMap (Esri Inc. 1999-2017). Land cover data were downloaded from the National Land Cover Database (Homer et al. 2015, U.S. Geological Survey 2014). The data were used with the extract by mask tool to remove areas covered by water from the understory cover model (Figure 34).

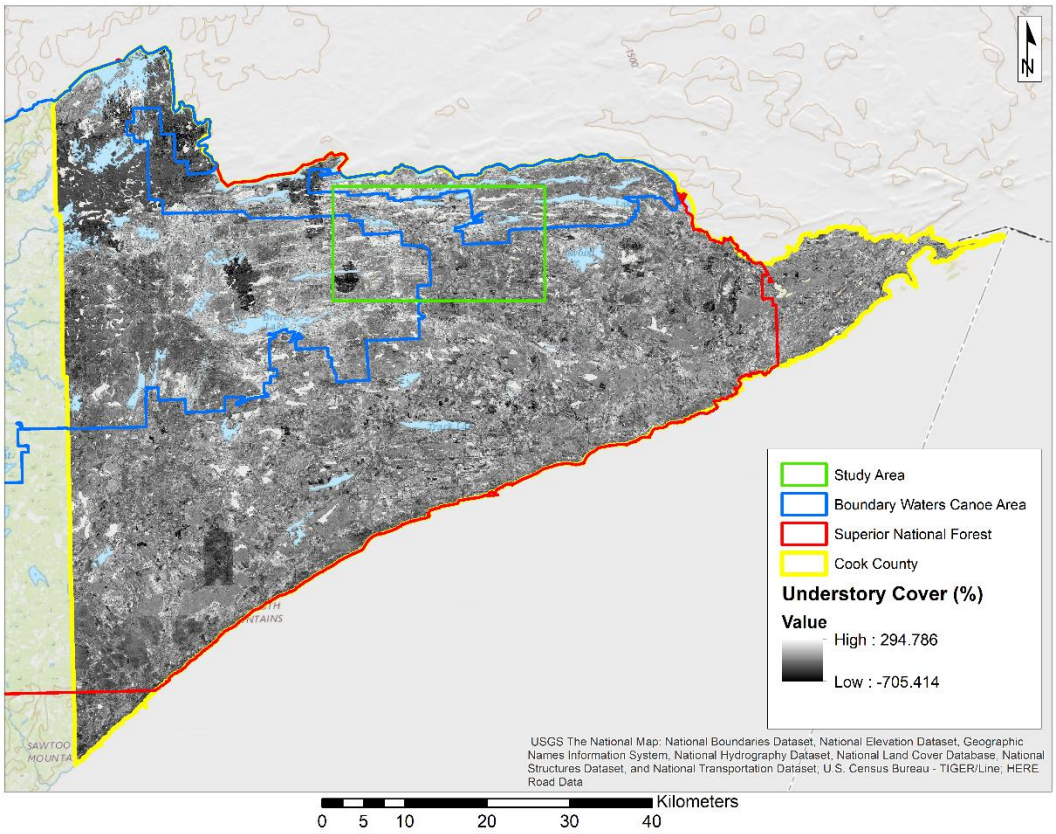

Figure 34. The "cleaned up" version of the understory cover model.

The trimmed model still contained values that were below the expected 0% understory cover, but many of the extreme values had been removed. The sample tool was again used to examine the cell values contained within the rasterized model. The histogram of the data appears to have a normal shape with a long-left tail (Figure 35).

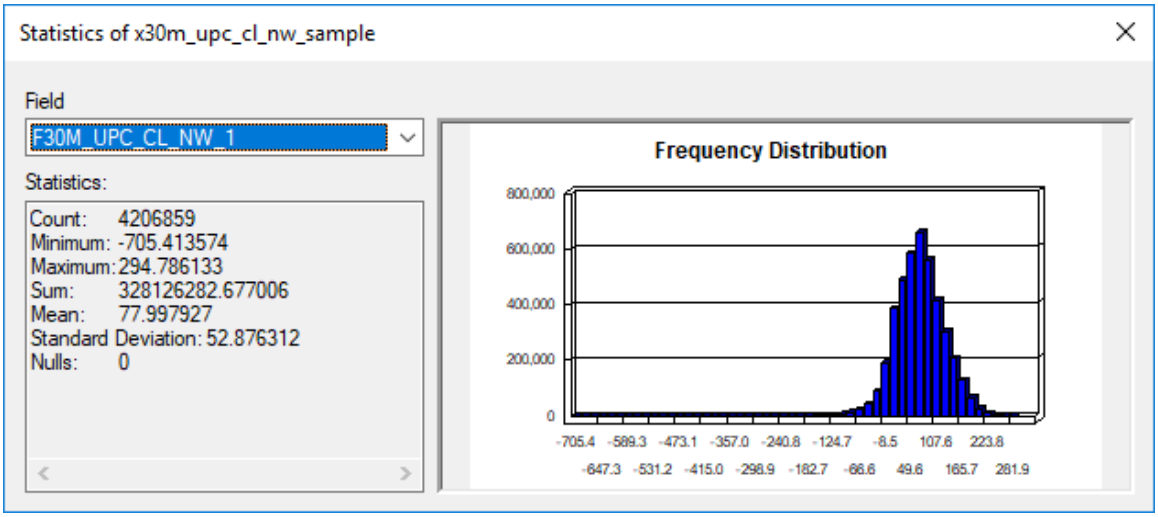

Figure 35. The pixel statistics and histogram for the understory cover model after trimming the spatial extent of the model.

Further analysis of the pixel values revealed that 242,403 of the 4,206,859 pixels (5.8%) were less than 0% understory cover, 42,785 (1.0%) were less than -50% understory cover, and 148 (0.004%) were less than -200% understory cover. The number of pixels at these break points were all less than that of the initial output of the model, although, proportionally they were higher. On the high side of the pixel values, 1,360,003 pixels (32.3%) were greater than 100% understory cover, 46,329 pixels (1.1%) were greater than 200% understory cover, and 703 pixels (0.02%) were greater than 250% understory cover. Again, the number of pixels was lower, but the proportion was higher.

In terms of the intercept value of the model (which was about 36.58% understory cover when converted from the canopy model intercept of about 7.36 meters) 886,642 of the 4,206,859 pixels, or 21.1%, fell below the intercept. In the initial output of the model, it was 1,025,218 out of 12,042,525 pixels (8.5%) that fell below the intercept value. In the initial output of the model, 7,454,920 pixels (61.9%) were equal to the

intercept, compared to 37,947 pixels (0.9%) in the cleaned-up version. The cleaned-up model had 3,282,270 pixels (78.0%) that were greater than the intercept, versus 3,562,387 pixels (29.6%) in the initial output of the understory cover model.

Rescaling of the trimmed model was attempted using a histogram of the pixel values and determining the percentage of pixels above and below selected threshold values. A new raster was generated using the following conditional statement in the raster calculator tool to set all values that were less than -140% understory cover (0.04% of the pixels) to -140% understory cover:

Con("30m_UPC_Cl_NW" < - 140, - 140, "30m_UPC_Cl_NW")

Subsequently, another raster was generated, which used the previously generated raster as the base, to set all of the values that were greater than 250% understory cover (0.02%) to 250% understory with the following conditional statement:

Con("UPC_30m_S4" > 250, 250, "UPC_30m_S4")

The product that resulted from these two equations can be seen in Figure 36. Visually, the raster appeared to be very similar to the cleaned-up understory cover model. The statistics showed that there was not a large change after the extreme high and low values had been adjusted (Figure 37). The mean of the cleaned-up raster was 77.997927% and the mean after the high and low values had been adjusted was 78.006862%, a difference of 0.0089325%. The sum of all pixel values in the cleaned-up raster was 328,126,282.677006% and the sum after the high and low values had been adjusted was 328,163,869.073643%, a difference of 37,586.39664% over 4,206,859 pixels. The standard deviation changed little, being 52.876312% before the adjustment of the high and low values and 52.824162% after the adjustment, a difference of -0.05215%.

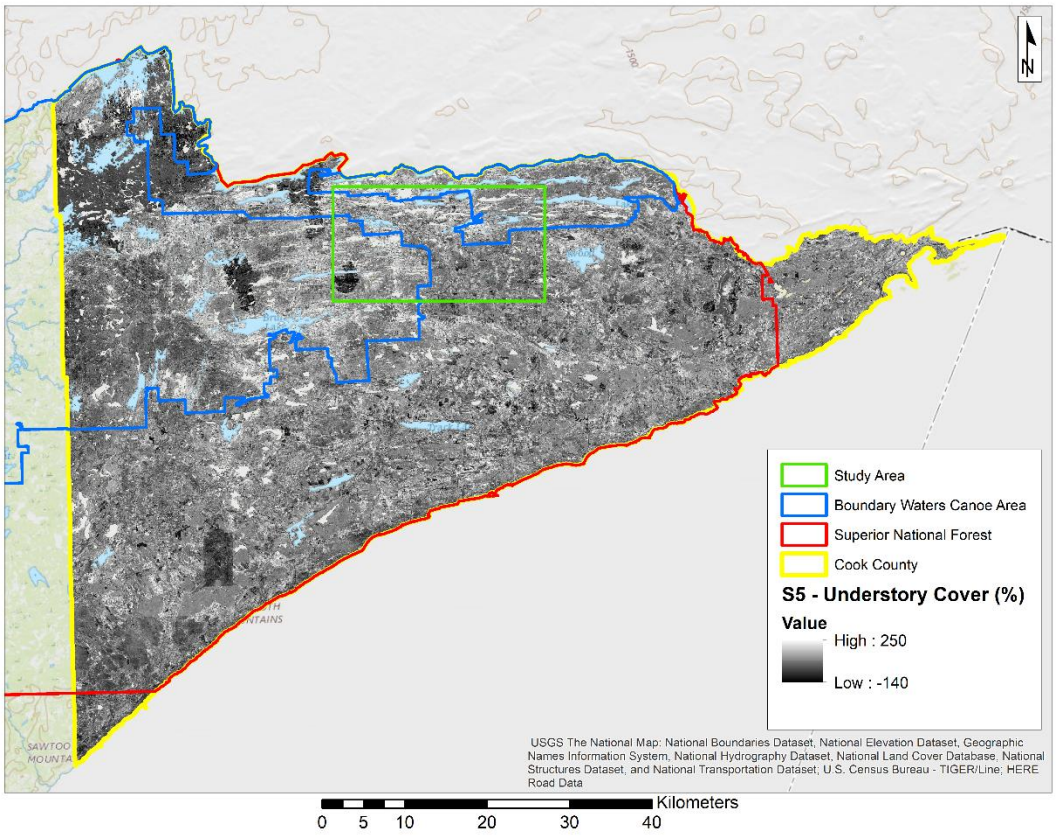

Figure 36. The raster product after adjusting the extreme high and low values.

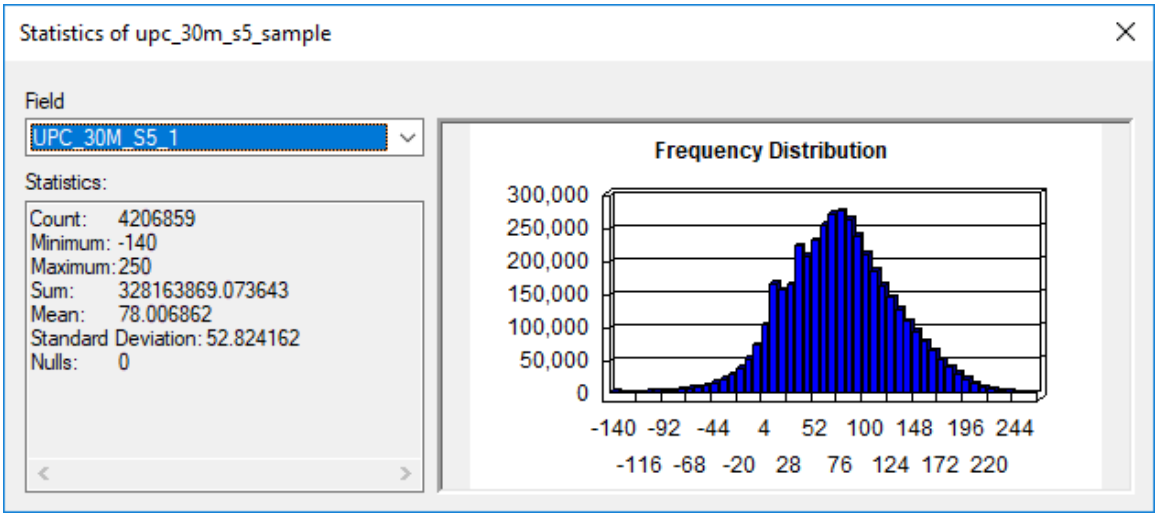

Figure 37. Statistics from the raster with the adjusted high and low values.

After the high and low values had been adjusted, the understory cover model was rescaled using the following formula from Esri Inc. (2016):

Rescaled grid = [(grid – Min value from grid) * (Max scale value - Min scale value) / (Max value from grid - Min value from grid)] + Min scale value

The actual equation that was used with the raster calculator tool was:

(("UPC_30m_S5" - (-140)) * (100 - 0) / (250 - (-140))) + 0

The resulting raster model can be seen in Figure 38 and the statistics can be seen in

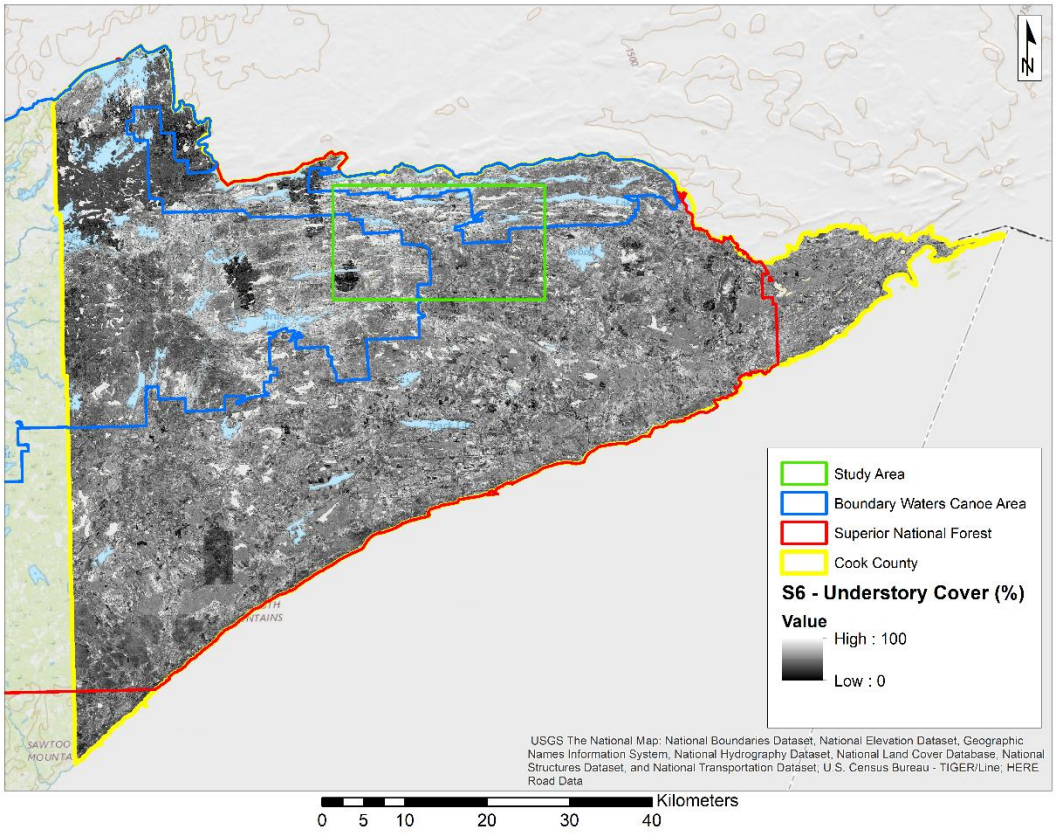

Figure 38. Rescaled understory cover model.

Figure 39. The rescaled understory cover model appeared to look similar to the cleaned-up version. The histogram of the pixel values also appeared to be similar in shape. The

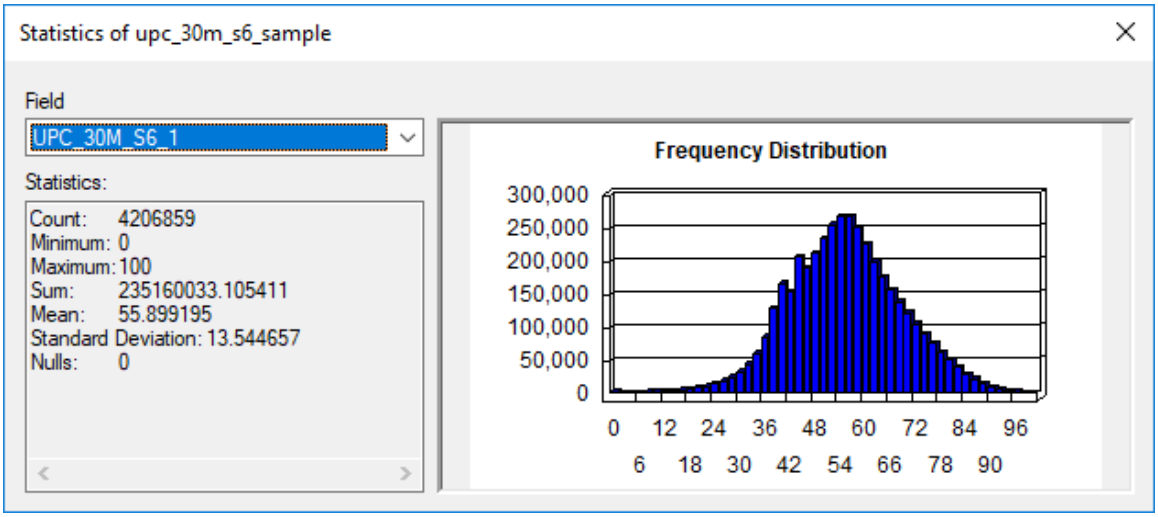

Figure 39. Statistics and histogram for the rescaled understory cover model.

mean decreased from around 78% to about 55.9% and the standard deviation dropped from around 52.8% to about 13.5%. The sum of the pixel values was also much lower.

It was decided that for the purposes of this investigation that using the rescaled model may not be appropriate. It was thought that in the future this product might be useful, however, when comparing data collected with the cover line intercept methodology it may not be.

Validation of the Understory Cover Model

Fieldwork to collect validation data occurred from May 16 – 26, 2017. Validation data were collected at 24 plots (Figure 9), using the cover line intercept method. All four transects were measured at all but one of those plots. Heavy rain forced the abandonment of the last three transects at plot 42. Photos were also taken at each plot, looking down each transect from the plot centers. Table 6 contains the summarized data for all of the validation plots.

Table 6. Plot totals of understory cover measurements for the validation plots. (U. Canopy = the understory canopy measured in meters, - indicates the transect was not measured.)

Plot 23	U. Canopy	% Cover	Plot 30	U. Canopy	% Cover
Trans 1	18.62 m	92.56%	Trans 1	27.80 m	138.16%
Trans 2	18.59 m	92.41%	Trans 2	14.87 m	73.93%
Trans 3	24.69 m	122.71%	Trans 3	11.61 m	57.72%
Trans 4	24.51 m	121.80%	Trans 4	14.51 m	72.11%
Mean	21.60 m	107.37%	Mean	17.20 m	85.48%
Plot 31	U. Canopy	% Cover	Plot 42	U. Canopy	% Cover
Trans 1	22.37 m	111.19%	Trans 1	31.79 m	158.01%
Trans 2	21.49 m	106.80%	Trans 2	-	-
Trans 3	20.42 m	101.50%	Trans 3	-	-
Trans 4	22.34 m	111.04%	Trans 4	-	-
Mean	21.66 m	107.63%	Mean	31.79 m	158.01%
Plot 48	U. Canopy	% Cover	Plot 56	U. Canopy	% Cover
Trans 1	34.23 m	170.12%	Trans 1	27.83 m	138.31%
Trans 2	25.97 m	129.07%	Trans 2	21.46 m	106.65%
Trans 3	26.30 m	130.74%	Trans 3	16.46 m	81.81%
Trans 4	28.93 m	143.77%	Trans 4	23.26 m	115.59%
Mean	28.86 m	143.42%	Mean	22.25 m	110.59%
Plot 57	U. Canopy	% Cover	Plot 59	U. Canopy	% Cover
Trans 1	20.42 m	101.50%	Trans 1	7.56 m	37.57%
Trans 2	22.01 m	109.38%	Trans 2	17.56 m	87.26%
Trans 3	26.37 m	131.04%	Trans 3	19.42 m	96.50%
Trans 4	23.59 m	117.25%	Trans 4	26.24 m	130.43%
Mean	23.10 m	114.79%	Mean	17.69 m	87.94%
Plot 110	U. Canopy	% Cover	Plot 139_ALT	U. Canopy	% Cover
Trans 1	32.83 m	163.16%	Trans 1	25.54 m	126.95%
Trans 2	25.69 m	127.71%	Trans 2	25.97 m	129.07%
Trans 3	34.35 m	170.73%	Trans 3	18.62 m	92.56%
Trans 4	32.28 m	160.43%	Trans 4	26.06 m	129.52%
Mean	31.29 m	155.51%	Mean	24.05 m	119.53%
Plot 154_W40	U. Canopy	% Cover	Plot 155	U. Canopy	% Cover
Trans 1	3.29 m	16.36%	Trans 1	14.33 m	71.20%
Trans 2	13.23 m	65.75%	Trans 2	19.35 m	96.20%
Trans 3	12.41 m	61.66%	Trans 3	21.92 m	108.92%
Trans 4	9.51 m	47.27%	Trans 4	20.18 m	100.29%
Mean	9.61 m	47.76%	Mean	18.94 m	94.15%

Table 6. Continued.

Plot 156	U. Canopy	% Cover	Plot 166	U. Canopy	% Cover
Trans 1	10.58 m	52.57%	Trans 1	19.69 m	97.86%
Trans 2	12.22 m	61.75%	Trans 2	20.48 m	101.80%
Trans 3	15.27 m	75.90%	Trans 3	30.08 m	149.52%
Trans 4	17.40 m	86.50%	Trans 4	17.74 m	88.17%
Mean	13.87 m	68.93%	Mean	22.00 m	109.34%
Plot 167	U. Canopy	% Cover	Plot 169_Extra	U. Canopy	% Cover
Trans 1	16.46 m	81.81%	Trans 1	4.69 m	23.33%
Trans 2	17.92 m	89.08%	Trans 2	13.08 m	64.99%
Trans 3	12.37 m	61.51%	Trans 3	4.54 m	22.57%
Trans 4	18.53 m	92.11%	Trans 4	6.64 m	33.03%
Mean	16.32 m	81.12%	Mean	7.24 m	35.98%
Plot 170	U. Canopy	% Cover	Plot 184	U. Canopy	% Cover
Trans 1	23.07 m	114.68%	Trans 1	21.73 m	108.01%
Trans 2	23.26 m	115.59%	Trans 2	21.15 m	105.13%
Trans 3	27.98 m	139.07%	Trans 3	20.60 m	102.41%
Trans 4	26.15 m	129.98%	Trans 4	19.54 m	97.11%
Mean	25.12 m	124.83%	Mean	20.76 m	103.17%
Plot 188	U. Canopy	% Cover	Plot 198	U. Canopy	% Cover
Trans 1	22.01 m	109.38%	Trans 1	16.09 m	79.99%
Trans 2	24.20 m	120.28%	Trans 2	26.37 m	131.04%
Trans 3	17.80 m	88.47%	Trans 3	23.99 m	119.22%
Trans 4	22.77 m	113.16%	Trans 4	19.20 m	95.44%
Mean	21.69 m	107.82%	Mean	21.41 m	106.42%
Plot 199	U. Canopy	% Cover	Plot 206_NW10	U. Canopy	% Cover
Trans 1	12.95 m	64.38%	Trans 1	9.63 m	47.87%
Trans 2	21.64 m	107.56%	Trans 2	12.86 m	63.93%
Trans 3	15.15 m	75.29%	Trans 3	14.81 m	73.62%
Trans 4	18.93 m	94.08%	Trans 4	17.25 m	85.74%
Mean	17.17 m	85.33%	Mean	13.64 m	67.79%
Plot 207	U. Canopy	% Cover	Plot 303	U. Canopy	% Cover
Trans 1	17.56 m	87.26%	Trans 1	9.14 m	45.45%
Trans 2	15.67 m	77.87%	Trans 2	10.73 m	53.32%
Trans 3	21.00 m	104.38%	Trans 3	11.80 m	58.63%
Trans 4	13.53 m	67.26%	Trans 4	6.37 m	31.66%
Mean	16.94 m	84.19%	Mean	9.51 m	47.27%

A total of 93 transects were measured in the 24 validation plots. The total amount of understory canopy measured for those transects was 1,799.42 meters, with a range

from 3.29 meters to 34.35 meters. The mean amount per transect was 19.35 meters. That equated to a minimum of 16.36% understory cover, a maximum of 170.73% understory cover, and a mean of 96.17% for the 93 transects.

For the 24 validation plots, the total of the mean understory canopy measured was 473.70 meters. The range for the plot means was narrower than for the transects, 7.24 meters to 31.79 meters, or 35.98% to 158.01% understory cover. The mean plot canopy measured was 19.74 meters, which was slightly greater than the mean measured transect canopy. The mean, in terms of understory cover, was 98.10% for the validation plots.

When the measured and predicted values (predicted values were strictly the understory cover pixel value for the pixel in which the plot was located) were compared for the validation plots, it appeared that some predictions were quite good, however, some missed the mark by quite a lot, particularly when the model was predicting negative understory cover (Table 7, Figure 40, upper right). The minimum difference (measured understory cover minus predicted understory cover) was -84.04 %, in plot 199, the maximum difference was 123.09%, in plot 169_Extra, and the mean difference was

Plot	Meas. U.C.	Pred. U.C.	Plot	Meas. U.C.	Pred. U.C.
23	107.37%	107.87%	30	85.48%	35.99%
31	107.63%	102.65%	42	158.01%	178.68%
48	143.42%	28.66%	56	110.59%	147.45%
57	114.79%	184.59%	59	87.94%	131.39%
110	155.51%	155.13%	139_ALT	119.53%	159.03%
154_W40	47.76%	-5.78%	155	94.15%	38.68%
156	68.93%	26.49%	166	109.34%	54.44%
167	81.12%	93.44%	169_Extra	35.98%	-87.11%
170	124.83%	82.33%	184	103.17%	94.42%
188	107.82%	68.69%	198	106.42%	170.57%
199	85.33%	169.37%	206	67.79%	21.52%
207	84.19%	105.70%	303	47.27%	-53.14%

Table 7. The measured and predicted understory values for the 24 validation plots. (Meas. U.C. = Measured Understory Cover, Pred. U.C. = Predicted Understory Cover).

14.30% (Figure 40, lower left). The minimum absolute difference between measured and predicted understory cover was 0.38%, in plot 23, the maximum absolute difference was 123.09%, in plot 169_Extra, and the mean absolute difference was 47.04% (Figure 40, lower right). The coefficient of determination for the measured versus predicted validation plots was 0.4698 (which equates to a correlation coefficient of 0.6854) (Figure 40, upper left) and was not nearly as good as the model's coefficient of multiple determination, 0.6504. It appeared that the understory cover model's predictions were all over the place when the lower left graph in Figure 40 was viewed. Overall, the model

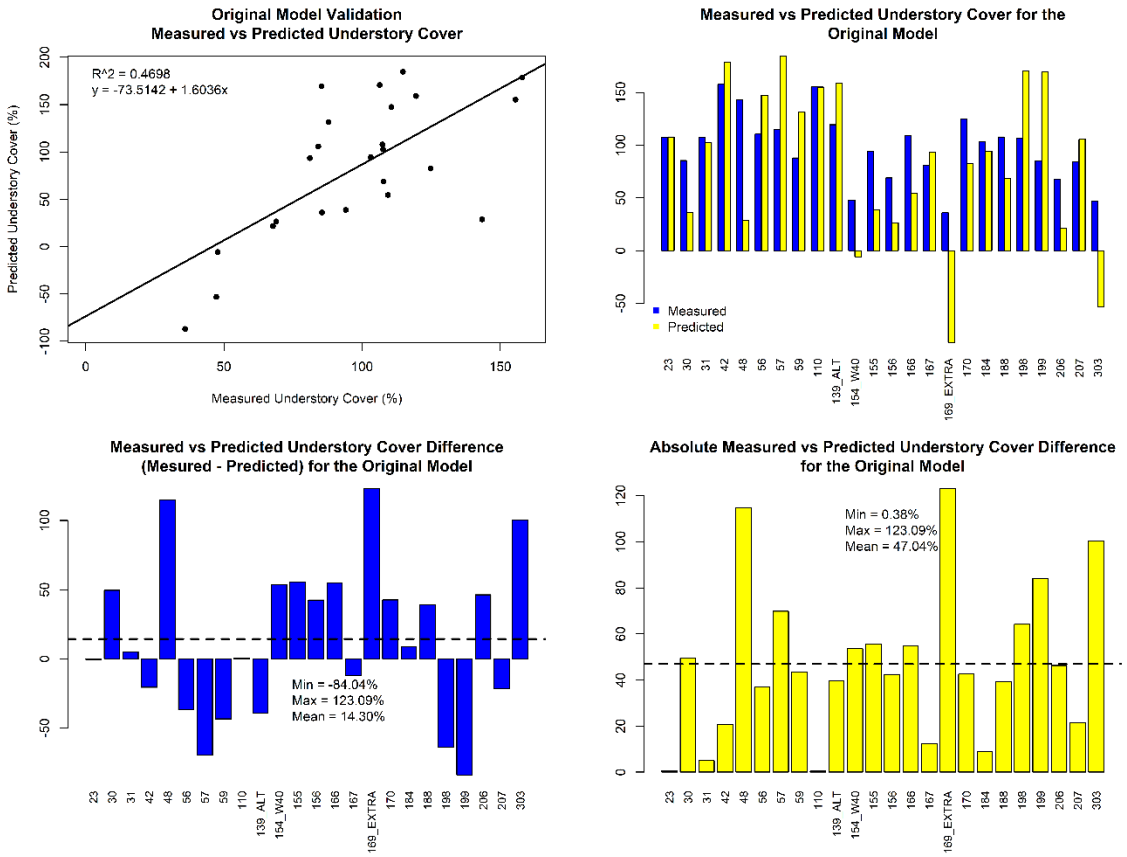

Figure 40. Graphs displaying the relationships between the measured and predicted understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory cover minus the predicted understory cover.

seemed to have a tendency to under predict the amount of understory cover (measured minus predicted), this was reflected in the mean difference between the measured and predicted values having been a positive 14.30%.

There may have been several reasons why the understory cover model and the validation did not have higher coefficients of determination. First of all, the lidar data itself was not particularly dense (less than 1 point per square meter), possibly having resulted in key points being missed. Better results could be possible with higher density lidar data.

Second, there was a considerable time lag between the airborne lidar collection date and when the field data was collected. The airborne lidar was collected in the spring of 2011, while the field data were collected in 2015, 2016, and 2017 field seasons. This time difference might have allowed additional vegetation growth. Disturbances may have also occurred during the time between the lidar collection and the fieldwork.

A third factor could be the lack of control plots. There were no phase 2 plots that had 0% understory cover. That certainly could have affected the model's intercept.

Furthermore, the plots in which only one transect, instead of all four, was measured could have influenced the results. Especially if the one transect was not representative of the plot. For instance, if a plot was relatively open in the northern portion, but full of understory trees in the rest of the plot, measuring only the northern transect would result in an underestimate of the total plot understory cover.

Along that line of thinking, it was likely that there were simply not enough plots measured with which to build the model. Also, the plots were mainly located near the Gunflint Trail or off one of the few "roads" connected to the Gunflint Trail. More plots

in a larger area would likely better represent the existing conditions within the study area. Future work could involve the collection of more plot data to improve the model's predictive power.

Another source of error could have been discrepancies in the way the data were collected by the different crews. Even the plot locations could be off. U.S. Forest Service Personnel did mark the center of each plot location with a wooden stake, however, if the stakes were placed with a handheld GPS unit, the location could have been off by 3 or 4 meters. Additionally, the lidar vendor reported that the horizontal accuracy of the lidar point cloud was "+/- 3.8 foot at the 95% confidence level," (Minnesota Department of Natural Resources 2014), which could have further added to the locational error.

Summary and Conclusion

An understory cover model was built by combining lidar metrics, derived from airborne lidar flown over Superior National Forest in spring of 2011, with field data collected using the cover line intercept method during the field seasons of 2015 and 2016. The understory cover model had a coefficient of determination of about 0.65. Field validation data were collected during the spring of 2017. Comparison of the validation data with predicted values from the model gave a correlation coefficient of about 0.69, with a mean difference of 14.3% and a mean absolute difference of 47.0%. While the understory cover model has room for improvement, the results appeared to indicate that it would be possible to map and measure understory vegetation using airborne lidar data.

COMPARISON OF THE UNDERSTORY COVER MODEL TO ANCILLARY DATA SETS

Introduction

Following development of the understory cover model, and prior to the validation work performed in 2017, the understory cover model was compared to some outside data sets. The first of the data sets was distributed by the U.S. Forest Service and was produced by two researchers from the University of Minnesota. Aaron Poznanovic and Michael J. Falkowski, had been working with the U.S. Forest Service in the same study area as this investigation (Poznanovic and Falkowski 2015). The second data set was the National Land Cover Database (Homer et al. 2015).

Material and Methods

University of Minnesota Data Sets

Poznanovic and Falkowski (2015) used 5 meter spatial resolution RapidEye imagery from June, 2013 and lidar data collected for the state of Minnesota in June, 2011, in combination with field data collected by the U.S. Forest Service, to map forest types, stand age, canopy bulk density, live crown base height, crown fuel base height, understory conifer cover, canopy cover, maximum canopy height, mean canopy height, and height of the 75th percentile in the same study area used for this investigation. The forest type (Figure 41), crown fuel base height (Figure 42), and live crown base height (Figure 43) products were investigated in this research. Crown fuel base height is the "height above the ground of the lowest live and/or dead fuels that have the ability to

move fire higher in the tree," and live crown base height is the "height of the lowest live branch whorl with live branches in two quadrants," (Lutes et al. 2006).

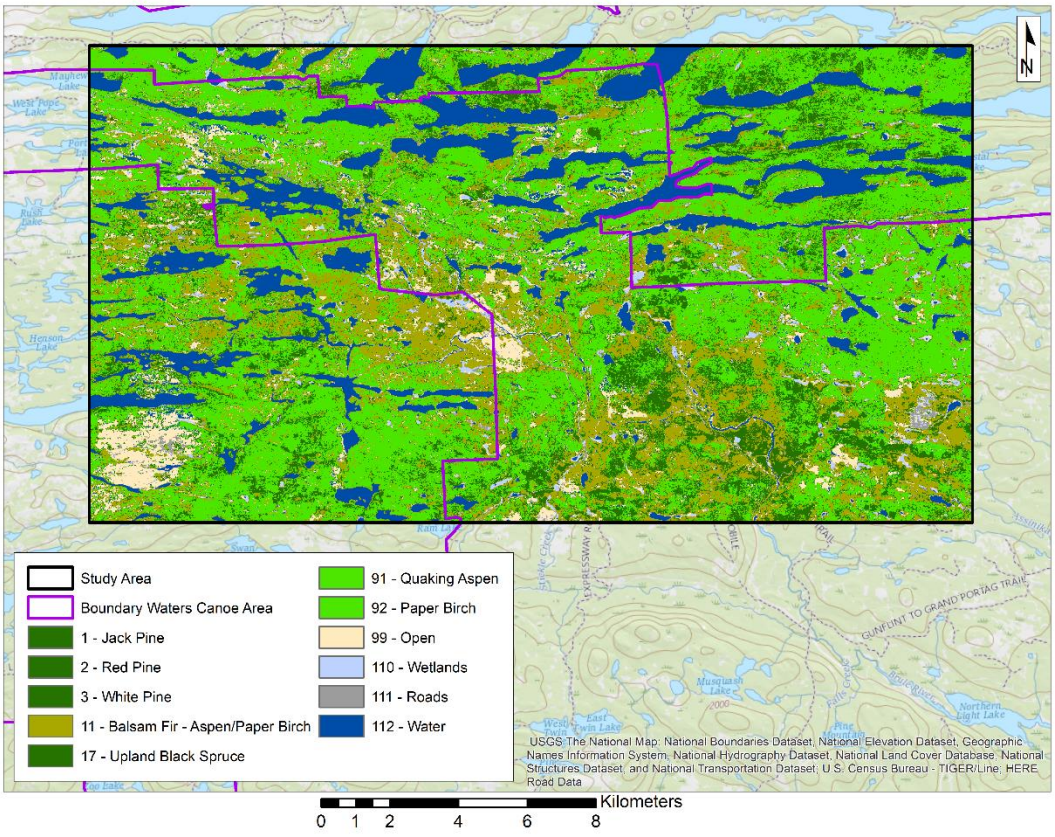

Figure 41. Map of forest types provided by the researchers from the University of Minnesota.

The forest type map was produced using a random forest machine-learning algorithm. The crown baseheight products were developed using nearest neighbor imputation. The overall accuracy of the forest type map was reported as 68.25%, with the highest individual accuracies being 61% for the red pine and 98% for the balsam fir – aspen/paper birch forest types. The coefficients of determination for the crown fuel base height and live crown base height products were 0.82 and 0.80, respectively.

In an attempt to identify a correlation between the mapped forest types and the

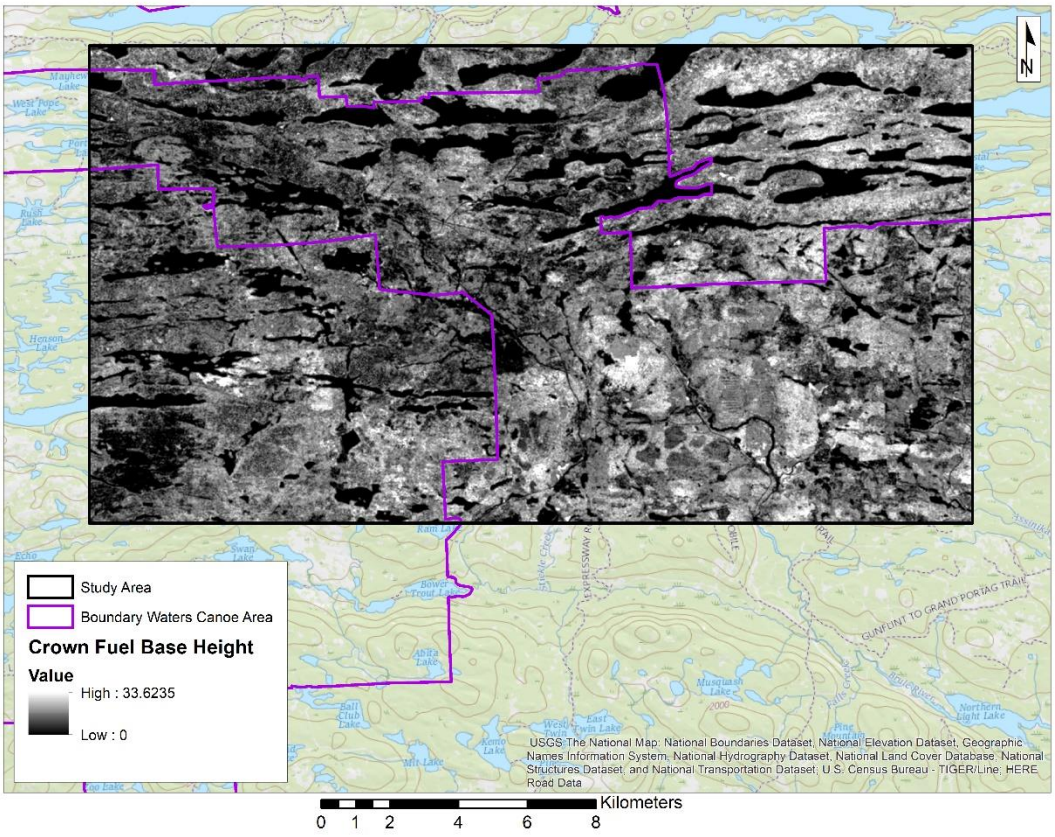

Figure 42. Map of crown fuel base height provided by the researchers from the University of Minnesota.

understory cover model, several binary rasters were generated from the original understory cover raster (the model prior to being cleaned-up) (Table 8). A binary raster is a raster layer that contains two values, usually listed as 0 and 1. The 1 value contains all of the pixels with the information that is of interest and the 0 value contains all of the pixels that are not of interest. For example, in the binary raster for the understory cover that is greater than 70%, all of the areas with an understory cover greater than 70% would be classified as 1 and all of the areas where the understory cover is less than or equal to 70% would be classified as 0 (Figure 44).

Conditional statements were used within the raster calculator tool to generate the

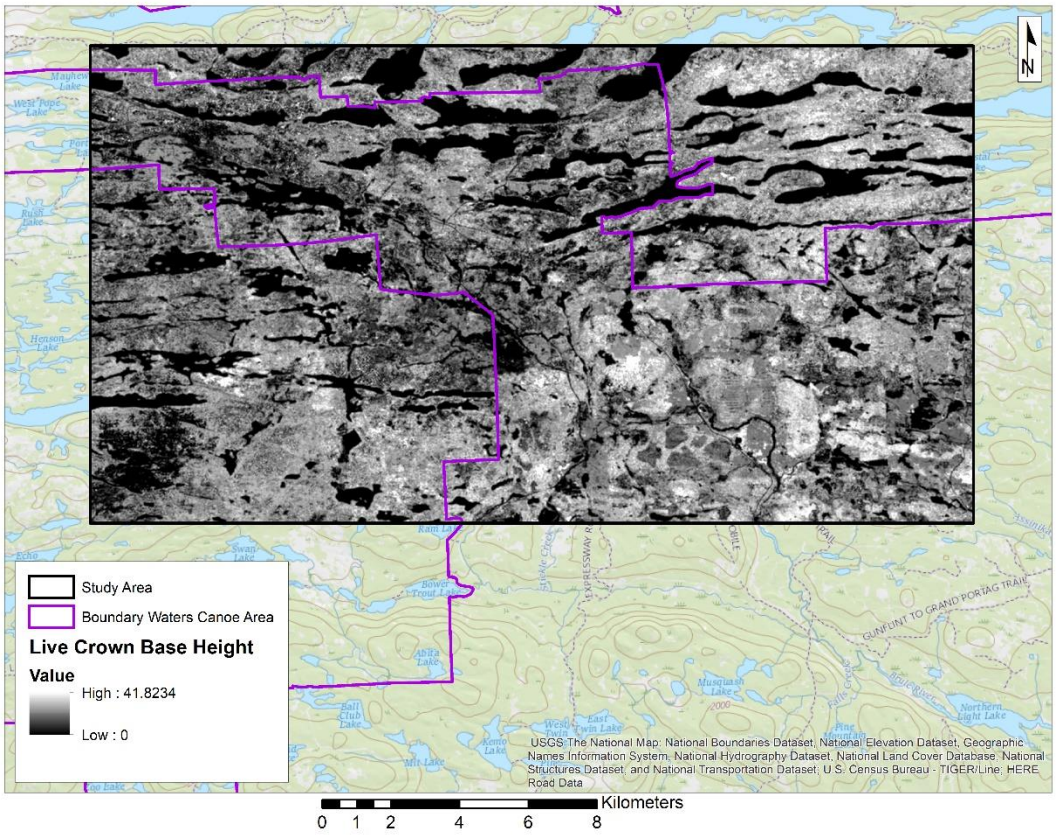

Figure 43. Map of live crown base height provided by the researchers from the University of Minnesota.

Understory Cover Binary Raster Layers		
Understory cover > -5,000%	Understory cover > 10%	Understory cover > 160%
Understory cover > -500%	Understory cover > 20%	Understory cover > 170%
Understory cover > -250%	Understory cover > 30%	Understory cover > 180%
Understory cover > -125%	Understory cover > 40%	Understory cover > 190%
Understory cover > -100%	Understory cover > 50%	Understory cover > 200%
Understory cover > -90%	Understory cover > 60%	Understory cover > 210%
Understory cover > -80%	Understory cover > 70%	Understory cover > 220%
Understory cover > -70%	Understory cover > 80%	Understory cover > 230%
Understory cover > -60%	Understory cover > 90%	Understory cover > 240%
Understory cover > -50%	Understory cover > 100%	Understory cover > 250%
Understory cover > -40%	Understory cover > 110%	Understory cover > 260%
Understory cover > -30%	Understory cover > 120%	Understory cover > 270%
Understory cover > -20%	Understory cover > 130%	Understory cover > 280%
Understory cover > -10%	Understory cover > 140%	Understory cover > 290%
Understory cover > 0%	Understory cover > 150%	

Table 8. List of the binary rasters generated for the comparison of the understory cover model with the ancillary data sets.

binary rasters. For instance, the equation used to generate the binary raster for understory cover greater than 70% was:

Con("SNF_30m_UCPC" > 70, 1, 0)

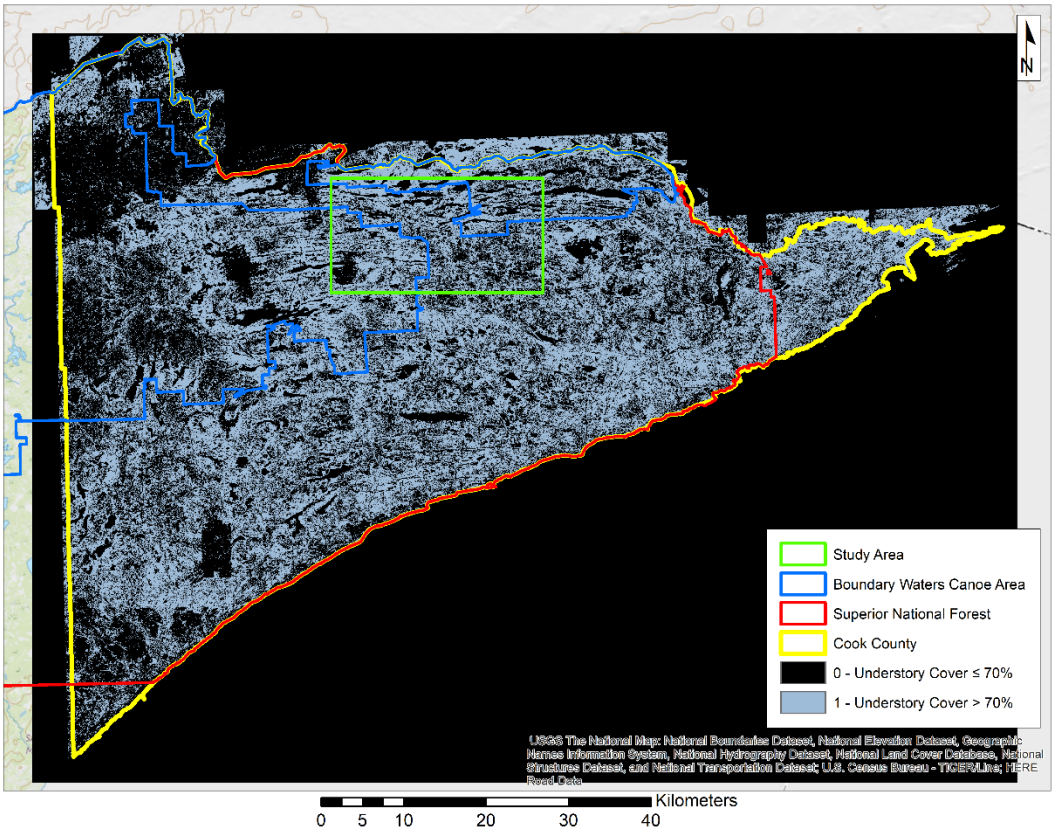

Figure 44. The binary raster displaying the areas where the understory cover is greater than 70%.

To link the various forest types mapped with the binary raster layers, ArcMap's (Esri Inc. 1999-2017) tabulate area tool was used to determine the area of each forest type that intersected each understory cover binary raster. The tabulate area tool calculates the area of a zone, defined in one raster, which is occupied by the class field specified in a second raster. The output is a table with information output as an area, not the number of pixels.

Both the crown fuel base height and the live crown base height data sets were also explored. The raster layers for both data sets were produced at 10-meter spatial resolution. ArcMap's (Esri Inc. 1999-2017) project raster tool was used to resample the layers to 30 meter resolution. The raster for the understory cover model was clipped down to the study area using the extraction by mask tool (Figure 45). The sample tool was used to generate a table containing the pixel values of the three raster layers. The understory cover model was then compared, individually, to each of the crown base height models, using linear regression.

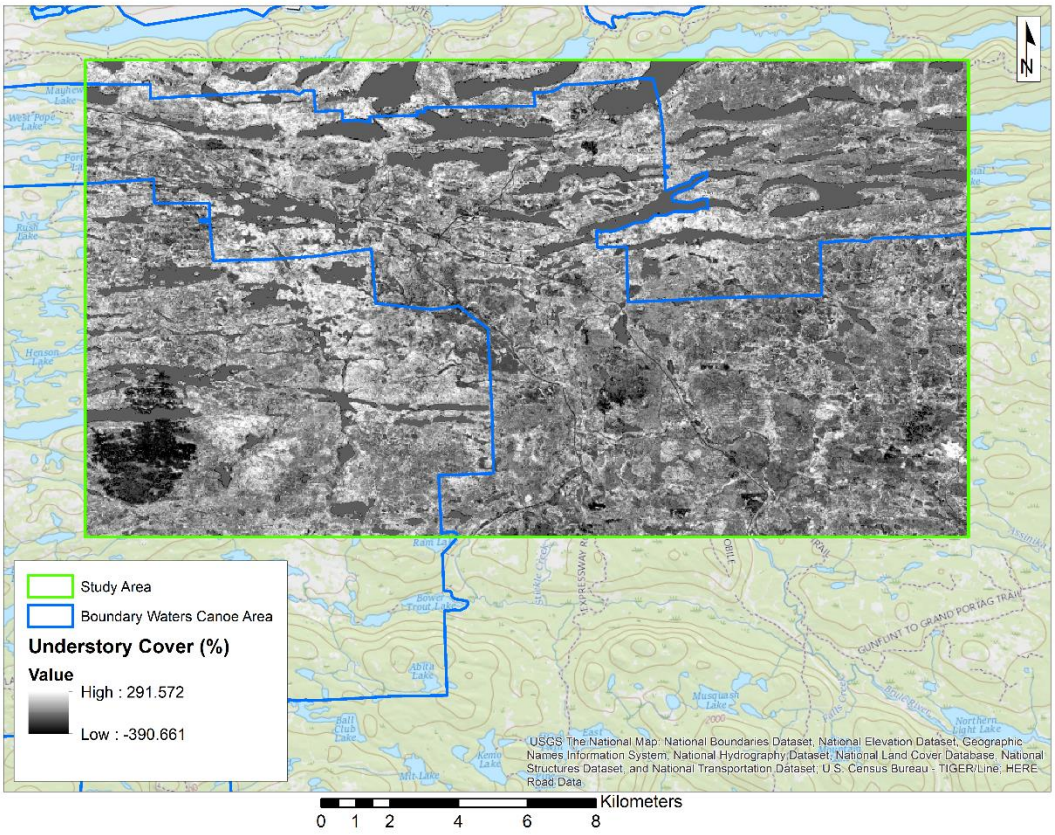

Figure 45. The understory cover model clipped to the study area.

Several binary rasters were also generated from the crown base height models. Two different approaches were used. In the first approach, the crown base height data sets were divided into a series of rasters showing areas containing data greater than the decided threshold heights. In the second approach, rasters were produced that put the data into height bins, for example crown fuel baseheight greater than 2 meters and less than or equal to 4 meters (Table 9). The tabulate area tool was used to compare the crown base height data sets to the forest type data.

Crown Fuel Base Height (meters)		Live Crown Base Height (meters)	
CFBH = 0	CFBH > 0	LCBH = 0	LCBH > 0
CFBH > 0 ≤ 2	CFBH > 2	LCBH > 0 ≤ 2	LCBH > 2
CFBH > 2 ≤ 4	CFBH > 4	LCBH > 2 ≤ 4	LCBH > 4
CFBH > 4 ≤ 6	CFBH > 6	LCBH > 4 ≤ 6	LCBH > 6
CFBH > 6 ≤ 8	CFBH > 8	LCBH > 6 ≤ 8	LCBH > 8
CFBH > 8 ≤ 10	CFBH > 10	LCBH > 8 ≤ 10	LCBH > 10
CFBH > 10 ≤ 12	CFBH > 12	LCBH > 10 ≤ 12	LCBH > 12
CFBH > 12 ≤ 14	CFBH > 14	LCBH > 12 ≤ 14	LCBH > 14
CFBH > 14 ≤ 16	CFBH > 16	LCBH > 14 ≤ 16	LCBH > 16
CFBH > 16 ≤ 18	CFBH > 18	LCBH > 16 ≤ 18	LCBH > 18
CFBH > 18 ≤ 20	CFBH > 20	LCBH > 18 ≤ 20	LCBH > 20
CFBH > 20 ≤ 22	CFBH > 22	LCBH > 20 ≤ 22	LCBH > 22
CFBH > 22 ≤ 24	CFBH > 24	LCBH > 22 ≤ 24	LCBH > 24
CFBH > 24 ≤ 26	CFBH > 26	LCBH > 24 ≤ 26	LCBH > 26
CFBH > 26 ≤ 28	CFBH > 28	LCBH > 26 ≤ 28	LCBH > 28
CFBH > 28 ≤ 30	CFBH > 30	LCBH > 28 ≤ 30	LCBH > 30
CFBH > 30 ≤ 32	CFBH > 32	LCBH > 30 ≤ 32	LCBH > 32
CFBH > 32 ≤ 34		LCBH > 32 ≤ 34	

Table 9. The binary rasters produced from the crown fuel base height (CFBH) and the live crown base height (LCBH) data sets. Heights are in meters.

The National Land Cover Database

The National Land Cover Database is a national land cover and land change data base that provides spatial data for the U.S. at a 30-meter resolution. Land cover and change products are available at 5-year intervals from 2001 to 2011 (Homer et al. 2015). The 2011 conterminous United States land cover data (Homer et al. 2015, U.S.

Geological Survey 2014) was downloaded for use in this research. The National Land Cover Database layer was clipped down to the study area using the extraction by mask tool (Figure 46). The National Land Cover Database layer was compared to the

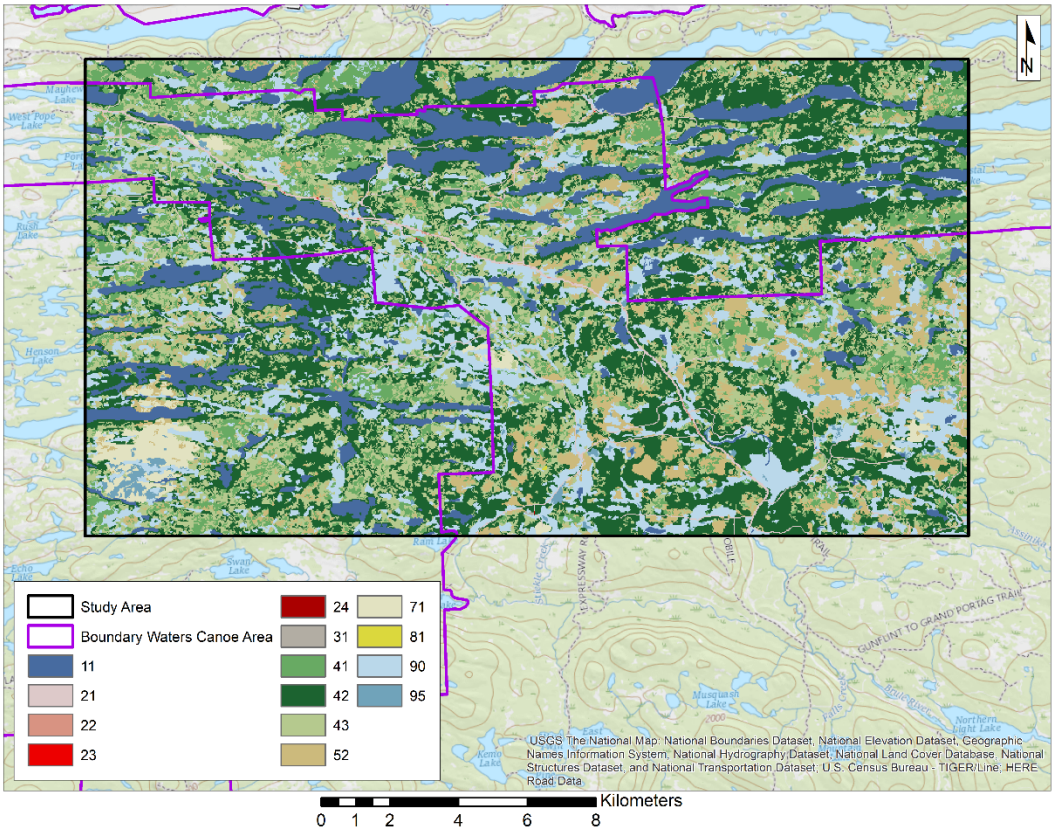

Figure 46. National Land Cover Database land cover map for the study area. (Land cover types: 11 = Open Water; 21 = Developed, Open Space; 22 = Developed, Low Intensity; 23 = Developed, Medium Intensity; 24 = Developed, High Intensity; 31 = Barren Land (Rock/Sand/Clay); 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 81 = Pasture/Hay; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands).

understory cover model using the same methodology as the comparison between the understory cover model and the University of Minnesota forest type data. The National Land Cover Database layer was also compared to the crown base height data sets using

the same methodology as the comparison of the crown base height data to the forest type map.

Results and Discussion

Understory Cover and Forest Types

Initial results from the comparison of the forest type map (Figure 41) to the understory cover model did not necessarily produce meaningful results. The results appeared to indicate that the forest types with the greatest amounts of area are the most important at ever break in understory cover (Figure 47). In an attempt to arrive at information that could be more meaningful, the different forest types were normalized by the proportion of total pixels that each forest type represented.

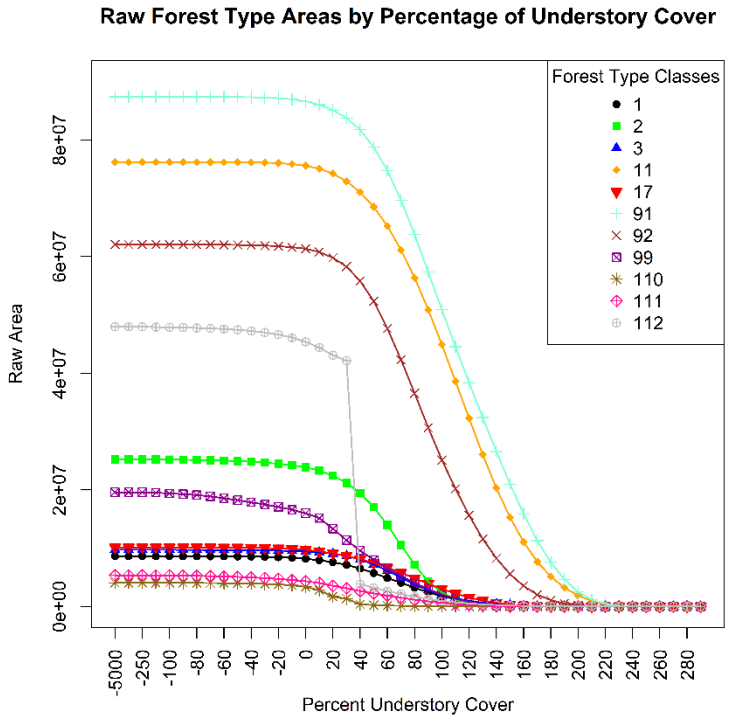

Figure 47. Comparison of the understory cover model to the forest type map. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands, 111 = roads, 112 = water).

The normalized results (Figure 48) indicated that as the percentage of understory cover increased, two classes of forest type stood out from the rest as representing the most area over a large portion the understory cover, the balsam fir – aspen/paper birch (class 11) and the quaking aspen (class 91) classes. The paper birch class appeared to be in a solid third in terms of normalized area as understory cover increased. The three pine

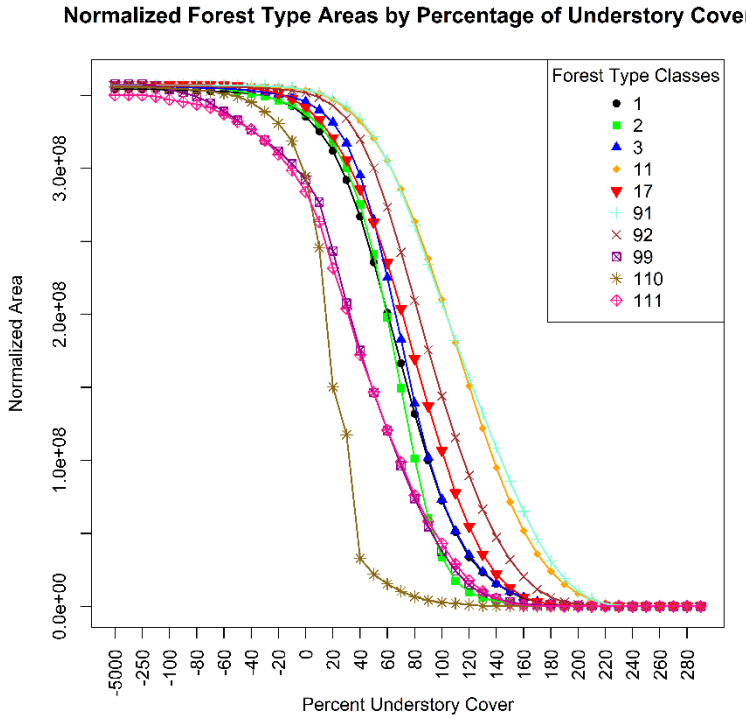

Figure 48. Comparison of the understory cover model to the forest type map, results have been normalized. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands, 111 = roads).

classes, jack, red, and white appeared to be about the same as the upland black spruce forest class in the lower portions of understory cover, but the upland black spruce started to have more area than the pine classes as understory cover approached 60%. It makes sense that areas that were classified as open, wetlands, or roads would occupy the least amount of normalized area at each level of understory cover, as these areas theoretically

should not have many trees. These results suggested that the balsam fir – aspen/paper birch, quaking aspen, and paper birch forests types could contribute the most to areas with high amounts of understory fuels.

Understory Cover and Crown Base Height

It was hypothesized that an inverse relationship would exist between the understory cover model and the crown base height models. Theoretically, if the height to which fuel is available for fire, crown base height, increases, the amount of fuel in the understory should decrease. Regression analysis indicated that there was almost no relationship between the models. The understory cover model had a correlation coefficient of 0.14 with the crown fuel baseheight model and 0.22 with the live crown baseheight model (Figures 49 and 50).

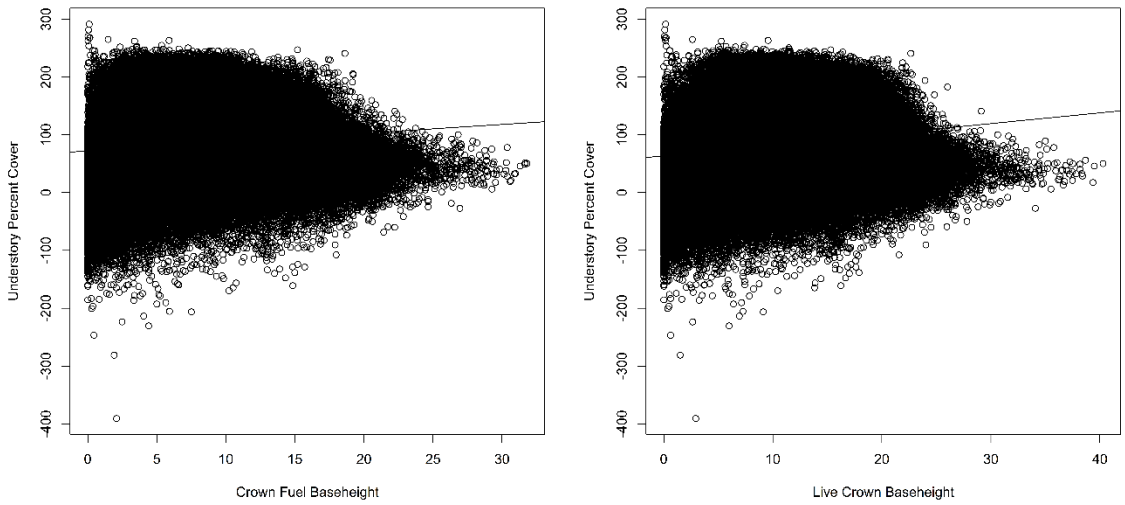

Figure 49. (Left) and Figure 50. (Right). Scatter plots showing the relationship between the understory cover model and the crown fuel base height (Figure 47) and live crown base height (Figure 48) models.

Forest Types and Crown Base Height

To explore the crown base height data sets further, the data sets were converted into two different series of binary rasters (Table 9). The first series was based on selected height threshold values and the second was based on dividing the models into height bins. These data sets were compared to the normalized forest type areas. When the forest type model and the crown fuel base height model were compared (Figures 51 and 52), it seemed that the three pine classes occupied the highest proportions of the area as the crown fuel base height increased. The order was white pine, followed by red pine, and

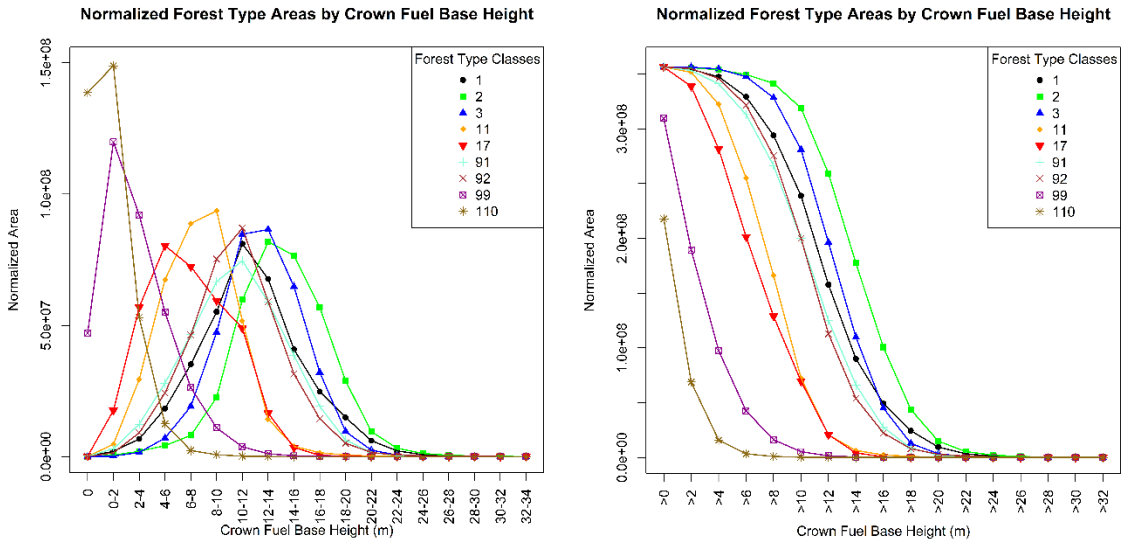

Figure 51. (Left). and Figure 52. (Right). Comparison of the crown fuel base height model to the forest type model. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands. Heights are in meters).

then jack pine. It would seem to make sense that white and red pine stands would have a higher crown fuel base height, as they are more adapted to surface fires.

Following the pines, came the quaking aspen and paper birch forest types, then balsam fir – aspen/paper birch, with upland black spruce last, in terms of the tree classes,

as the crown fuel base height increased. Again, these results seemed to be logical, as the deciduous species do not, typically, have a lot of lower limbs. Adding balsam fir in with aspen and birch should lower the crown fuel base height as balsam fir typically have more continuous canopies. Spruce usually have continuous canopies as well, which was indicated by upland black spruce stands having the lowest crown fuel base heights.

The results for the comparison between the forest types and the live crown base height model appeared to be very similar, with the main difference being that the jack pine forest type results were closer to those of the quaking aspen and paper birch forest types (Figures 53 and 54).

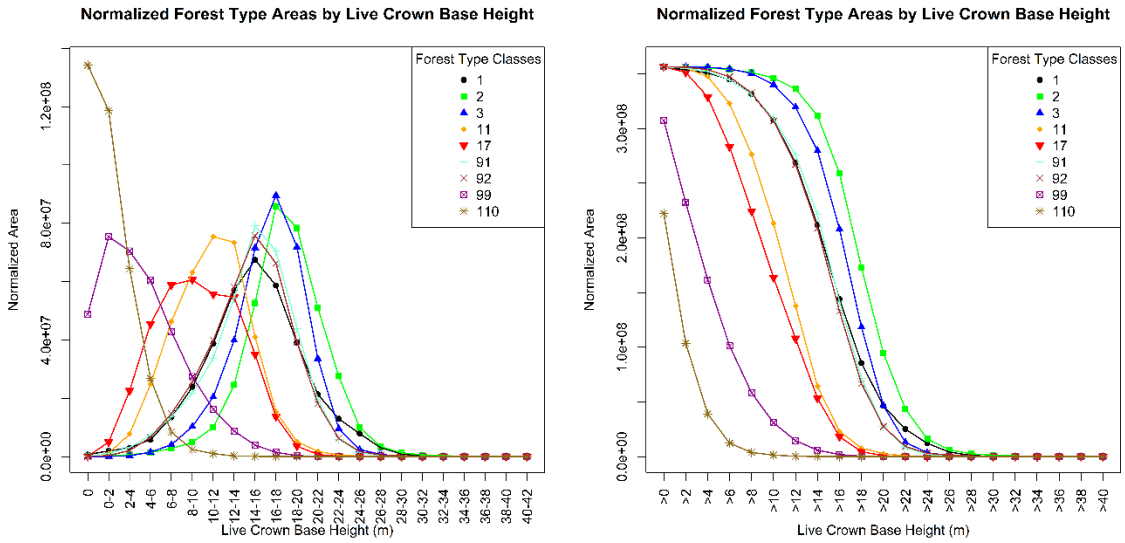

Figure 53. (Left) and Figure 54. (Right). Comparison of the live crown base height model to the forest type model. (Forest Type Class: 1 = jack pine, 2 = red pine, 3 = white pine, 11 = balsam fir - aspen/paper birch, 17 = upland black spruce, 91 = quaking aspen, 92 = paper birch, 99 = open, 110 = wetlands. Heights are in meters).

Understory Cover and the National Land Cover Database

In addition to the forest type map, another land cover data set was compared to the understory cover model, the National Land Cover Database (Homer et al. 2015, U.S.

Geological Survey 2014). As with the forest type map, the results from the initial comparison with the understory cover model did not have much meaning (Figure 55). However, when normalized (Figure 56), patterns became more apparent. As understory cover increased, the deciduous forest class seemed to make up the highest proportion of the area. Next came the mixed forest class. Evergreen forest appeared to be much more important at lower levels of understory cover than the shrub/scrub class, but as understory cover increased past about 60% - 70%, the shrub/scrub class was higher in proportion. The woody wetlands class, followed distantly by the emergent herbaceous wetlands class, and even more distantly by the grassland/herbaceous class rounded out the vegetation classes.

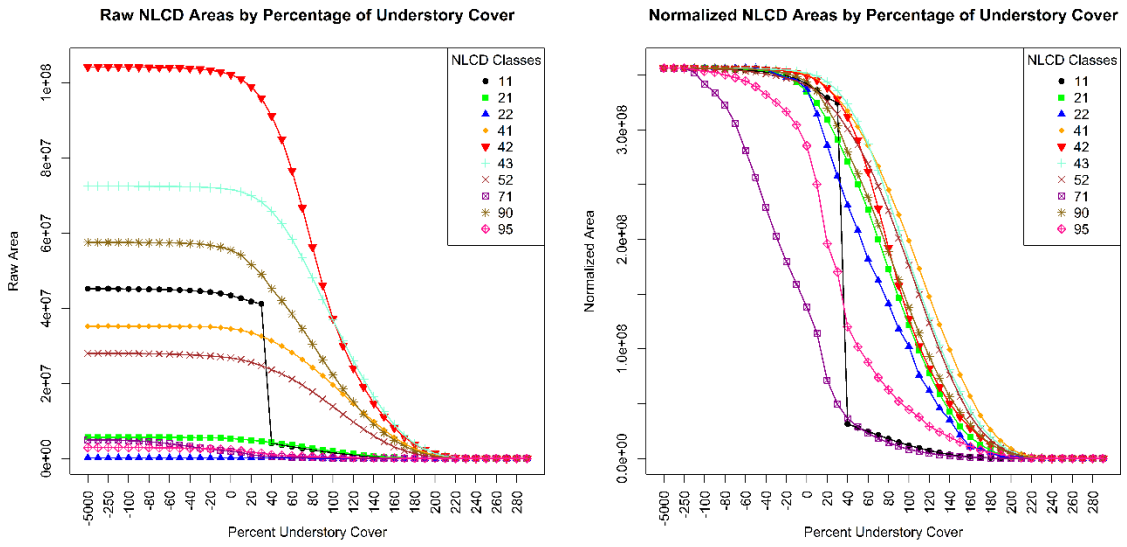

Figure 55. (Left). and Figure 56. (Right). Comparison of the understory cover model to 2011 land cover data from the National Land Cover Database. Class areas have been normalized in Figure 54. (Land cover classes: 11 = Open Water; 21 = Developed, Open Space; 22 = Developed, Low Intensity; 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands).

These results appeared to be in agreement with the results from the forest type comparison. In the forest type comparison, the quaking aspen forest type (deciduous)

appeared to have the most proportional area as the understory cover amounts increased, followed by the balsam fir – aspen/paper birch forest type (mixed coniferous and deciduous) and the paper birch forest type (again, deciduous). The upland black spruce forest type and the three pine forest types (all coniferous) trailed the paper birch class. It was difficult, however, to make comparisons with the shrub/scrub class, as a similar class was not specifically mapped by Poznanovic and Falkowski (2015).

The National Land Cover Database and Crown Base Height

The National Land Cover Database data (Homer et al. 2015, U.S. Geological Survey 2014) were also compared to the crown base height binary rasters. The results of comparison with the crown fuel base height model (Figures 57 and 58) indicated that the evergreen class had the highest proportion of the area as crown fuel base height increased, followed by the mixed forest class, and the deciduous forest class. The

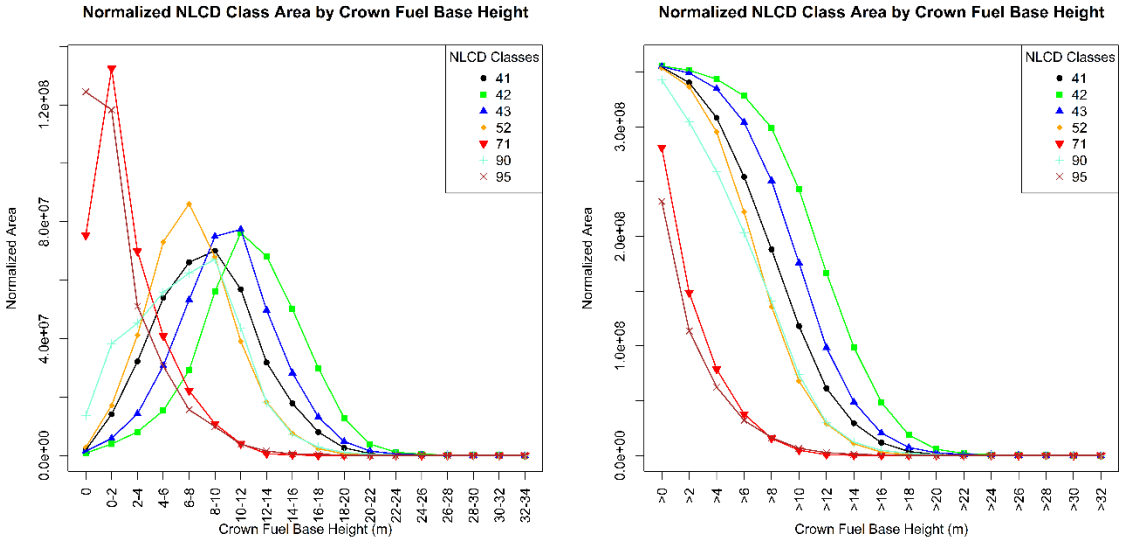

Figure 57. (Left). and Figure 58. (Right). Comparison of the crown fuel base height model to land cover data from the National Land Cover Database. (Land cover classes: 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands).

shrub/scrub class was next, with woody wetlands close behind. The grassland/herbaceous and emergent herbaceous wetlands classes dominated the low end of crown fuel baseheight, but dropped off precipitously.

The graphs looked very similar for the comparison of the land cover data with the live crown base height model (Figures 59 and 60). One very curious result was the peaks of the shrub/scrub class at crown fuel base height greater than 6 meters and less than or equal to 8 meters (Figure 57) and live crown base height greater than 10 meters and less than or equal to 12 meters (Figure 59).

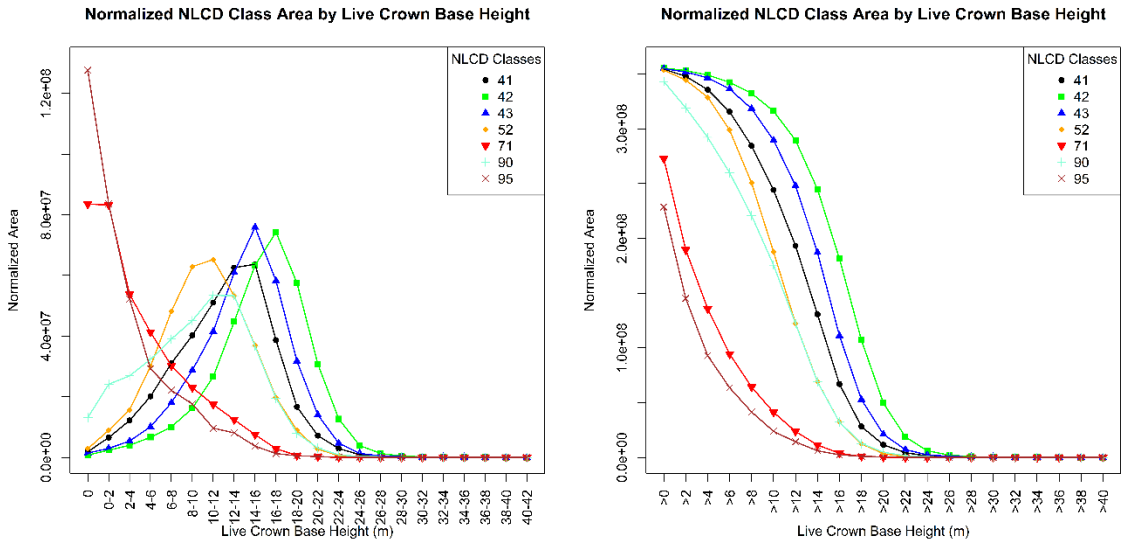

Figure 59. (Left). and Figure 60. (Right). Comparison of the live crown base height model to land cover data from the National Land Cover Database. (Land cover classes: 41 = Deciduous Forest; 42 = Evergreen Forest; 43 = Mixed Forest; 52 = Shrub/Scrub; 71 = Grassland/Herbaceous; 90 = Woody Wetlands; 95 = Emergent Herbaceous Wetlands).

Generally speaking, the results again appeared to agree with forest type model comparison. In the forest type comparison, the pine classes had the highest proportions of the area at higher levels of crown base height, corresponding well with the evergreen class. The results were a little mixed after that. For the forest type comparison, the

deciduous species aspen and birch followed the pines, not the mixed forest type of balsam fir – aspen/paper birch.

Summary and Conclusion

The understory cover model was compared to ancillary data sets, primarily through the use of binary rasters. The ancillary data sources include three data sets produced by researchers from the University of Minnesota, a forest type map and two different crown base height models (Poznanovic and Falkowski 2015). The understory cover model was also compared to land cover data from the National Land Cover Database (Homer et al. 2015, U.S. Geological Survey 2014).

Additionally, both land cover data sets, the forest type map (Poznanovic and Falkowski 2015) and the National Land Cover Database data (Homer et al. 2015, U.S. Geological Survey 2014), were compared to the crown fuel base height and live crown base height models of Poznanovic and Falkowski (2015).

Results from the comparison of the understory cover model with the forest type data (Poznanovic and Falkowski 2015) indicated that there was more area of the balsam fir – aspen/paper birch and quaking aspen forest types as the understory cover increased. The comparison of the understory cover model with the National Land Cover Database data set (Homer et al. 2015, U.S. Geological Survey 2014) seemed to corroborate these findings. No relationship was found between the understory cover model and the crown base height data sets produced by Poznanovic and Falkowski (2015).

Comparison of the forest type data with the crown base height models (Poznanovic and Falkowski 2015) indicated that pine stands have the highest

proportional amount of area as crown base height increased, followed by deciduous stands, and then mixed stands. The comparison between the crown base height models (Poznanovic and Falkowski 2015) and the National Land Cover Database land cover data (Homer et al. 2015, U.S. Geological Survey 2014) seemed to mostly be in agreement.

BALSAM FIR UNDERSTORY COVER MODEL DEVELOPMENT

Introduction

Extracting the balsam fir component of the understory vegetation from the lidar data was a highly desired goal of this investigation. Unfortunately, the lidar data do not contain spectral information. Also, after having looked at numerous point clouds, field photographs, and plot data, it was decided that the effort to extract balsam fir data directly from the lidar was impractical.

A collaboration with Dr. Peter Wolter, from Iowa State University, was established through the U.S. Forest Service. Dr. Wolter had worked in the Superior National Forest Region previously (see Sturtevant et al. 2012, Wolter et al. 2008, Wolter, Townsend, and Sturtevant 2009, Wolter and Townsend 2011). Wolter was also focusing on the same study area as this project. Dr. Wolter's goal was to map the basal area of several tree species, with future work to focus on mapping canopy bulk density (Wolter 2016b). Basal area is the amount of cross sectional area occupied by tree stems, measured at breast height, in a given unit of area, for example square meters per hectare. Dr. Wolter shared the data for his work in Superior National Forest (Wolter 2016a).

Materials and Methods

Dr. Wolter collected basal area field data at several variable radius plots using a metric basal area factor prism. Plot radii were determined by the theoretical intersection of a 23º cone, looking upward from the plot center, and the canopy height. Wolter correlated the field data with five Landsat-8 scenes from 2014 and 2015, three from winter dates, one summer scene, and one fall scene. The winter scenes were incorporated

to help reduce the occlusion of understory conifers. Out of this effort came models for total basal area, conifer basal area, hardwood basal area, balsam fir basal area, white spruce basal area, black spruce basal area, tamarack basal area, cedar basal area, white pine basal area, and red pine basal area. Those models had coefficients of determination of 0.83, 0.93, 0.72, 0.68, 0.32, 0.77, 0.78, 0.32, 0.21, and 0.84 respectively and root mean square errors (in square meters per hectare) of 3.04, 2.96, 1.57, 2.31, 1.73, 2.83, 0.46, 1.36, 0.71, and 4.28 respectively. The models were mapped as raster data sets with basal area in square meters per 30-meter pixel (Wolter 2016b).

The balsam fir basal area raster data (Wolter 2016b) (Figure 61) was explored through the use of binary rasters . The balsam fir basal area data were classified into five

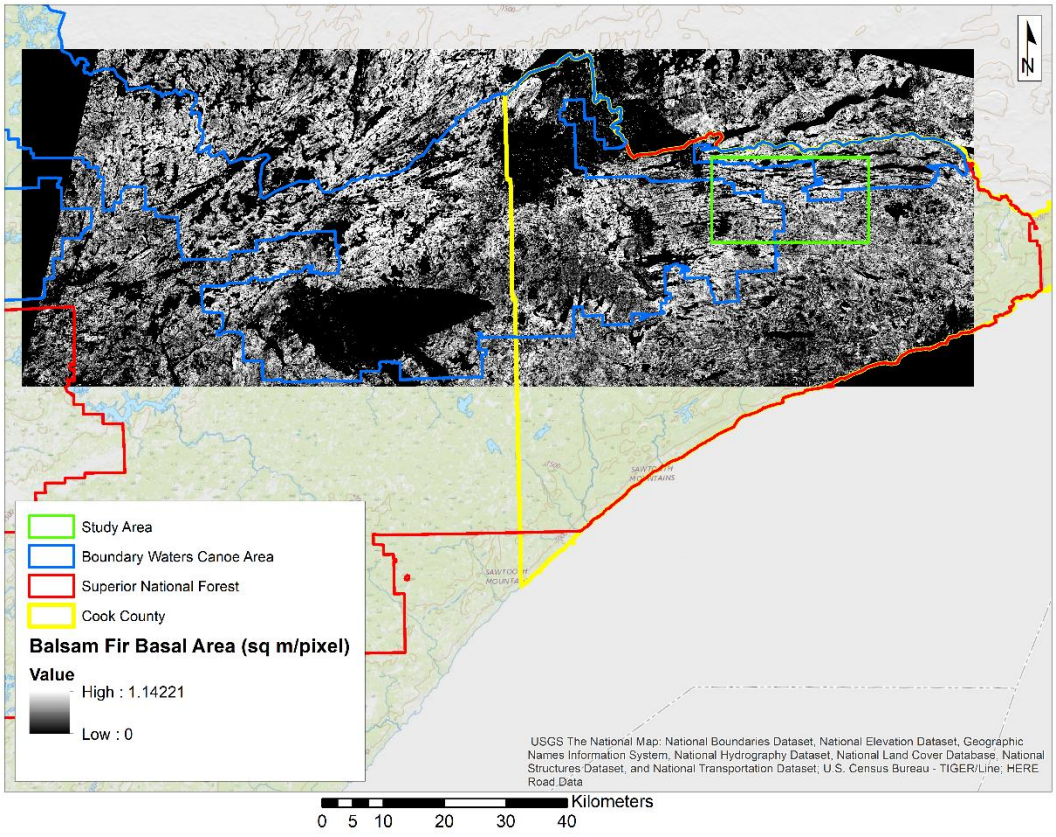

Figure 61. The balsam fir basal area, measured in square meters per pixel (Wolter 2016b).

classes using a natural breaks (Jenks) classification. Each of the five classes was then converted into a binary raster using the raster calculator tool and clipped down to the same extent as the understory cover binary rasters using the extraction by mask tool. The proportional amounts of each class of balsam fir basal area were calculated from the binary rasters. These data were then used to normalize each balsam fir basal area class. Both raw and normalized balsam fir basal area class data were compared to understory cover model binary rasters via the tabulate area tool.

A proportion of balsam fir raster was generated by using raster calculator to divide the balsam fir basal area raster (Figure 61) by the total (all species) basal area raster (Wolter 2016b) (Figure 62). A modified version of that raster was then multiplied

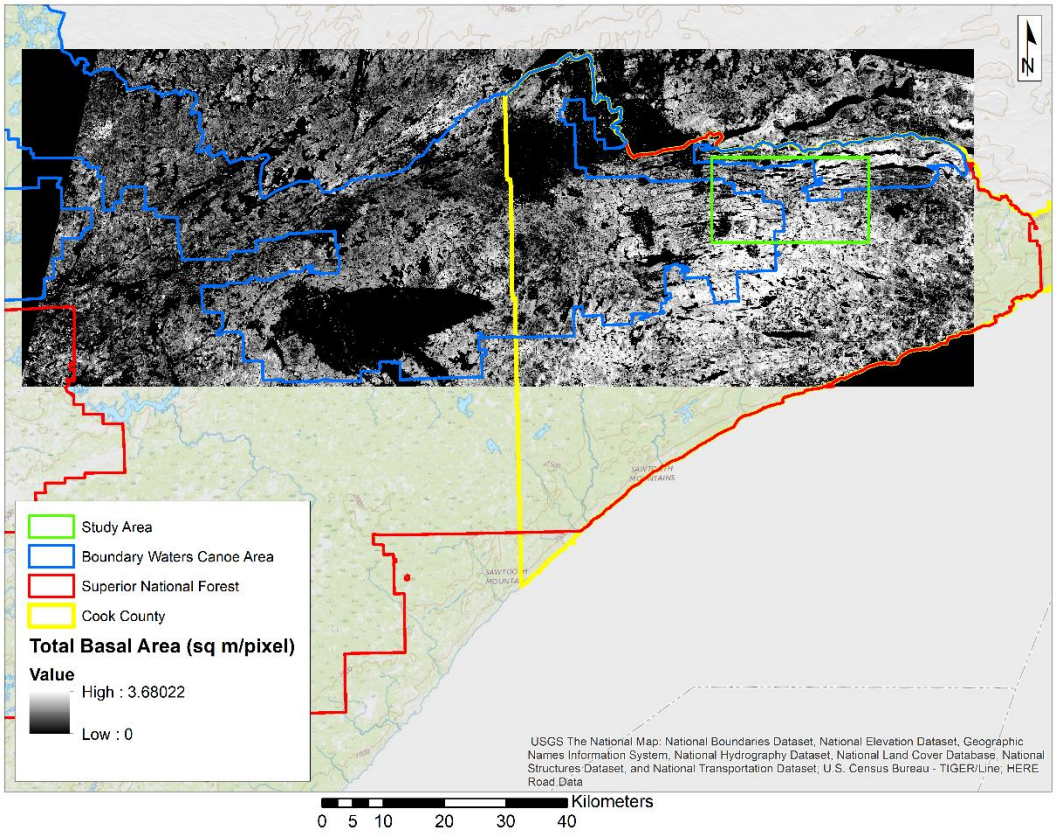

Figure 62. Total basal area spatial model (Wolter 2016b). Basal area is in square meters per pixel.

by the original, non-rescaled, understory cover (Figure 34) spatial model and the rescaled understory cover model (Figure 38) to calculate a theoretical amount of balsam fir understory cover. Both versions of balsam fir understory cover raster were validated using the cover line intercept field data collected in May of 2017.

Results and Discussion

The balsam fir basal area data (Wolter 2016b) were broken into classes based upon the natural breaks (Jenks) classification (Table 10). Those five classes were then used to generate a series of binary raster for each class (Figure 63). Comparison of the

Balsam Fir Basal Area Classes	**Basal Area in m^2/pixel**
Class 1	0 - 0.134377858
Class 2	0.134377858 - 0.309069074
Class 3	0.309069074 - 0.456884718
Class 4	0.456884718 - 0.613658886
Class 5	0.613658886 - 1.142211795

Table 10. Natural breaks (Jenks) classes generated from the balsam fir basal area raster.

balsam fir binary rasters with the understory cover binary rasters produced some interesting results. Both raw and normalized balsam fir basal area data were compared to the binary understory cover rasters and in both cases, there appeared to be a large break at 30% understory cover (Figures 64 and 65). From understory cover greater than 40% to about understory cover greater than 230% class 5 balsam fir basal area appeared to be the most dominant class in terms of both raw and normalized area, with class 1 occupying the greatest amount of raw area at lower understory covers. However, after normalization, it appeared that class 1 balsam fir basal area occupied the least amount of area, proportionally. The raw results indicated that there was not a large of amount of area in

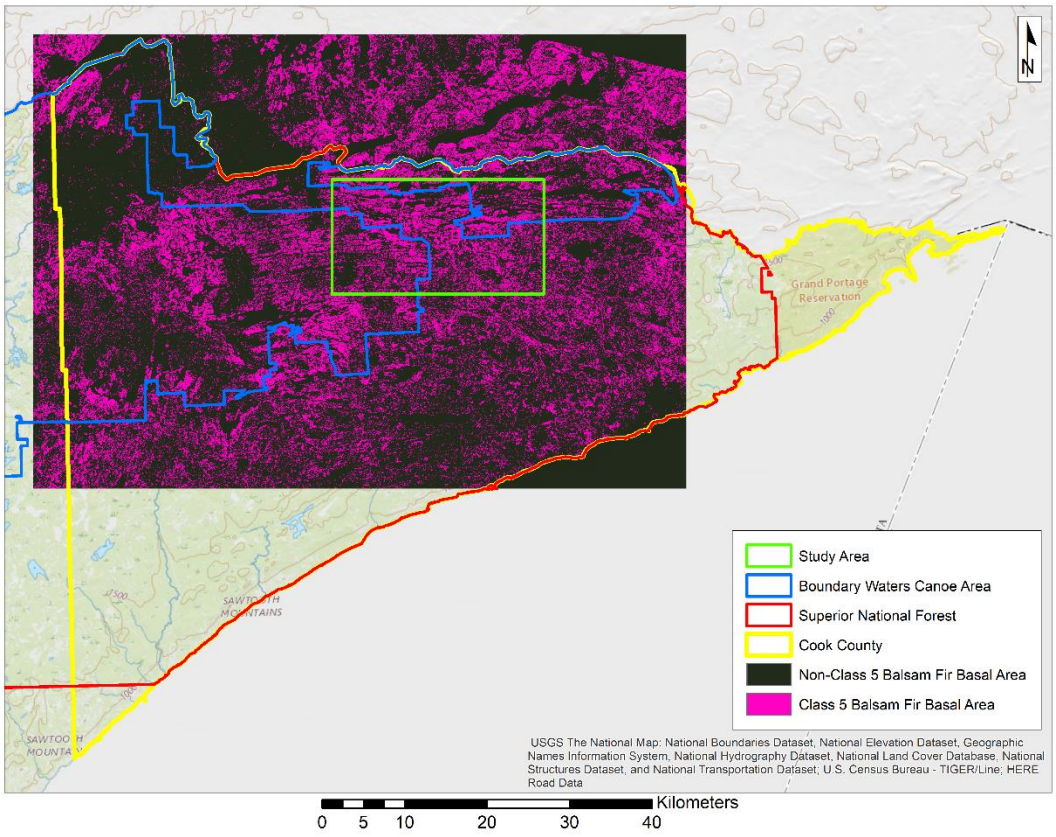

Figure 63. The class 5 balsam fir binary raster.

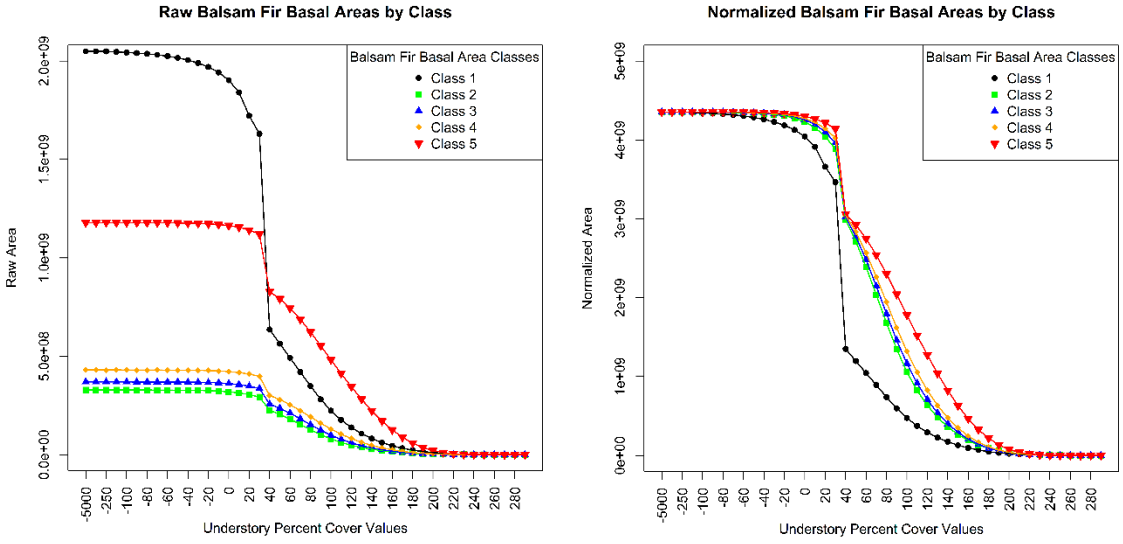

Figure 64. (Left) and Figure 65. (Right). Comparison of the raw (Figure 62) and normalized (Figure 63) amount of area in each of the five-balsam fir basal area classes.

classes 2 through 4, but that within those three classes, class 4 had the largest amount of area, followed by class 3, and finally class 2, although there was not a large amount of difference between the three. When the data were normalized the gap between classes 2, 3, and 4 appeared to narrow, but the order remained the same. Classes 2, 3, and 4 were also proportionally much higher in the normalized area than class 1, but still distinctly lower than class 5. These results suggested that there was a correlation between high amounts of balsam fir basal area and moderate to high amounts of understory cover.

A proportion of balsam fir raster was created dividing the balsam fir basal area raster (Figure 61) by the total (all species) basal area raster (Figure 62) (Wolter 2016b). The resulting raster had an interesting range, 0 to 163,787 (theoretically the proportion should not exceed 1). The sample tool was used to examine the individual pixel values. It was determined that 675,342 of the 5,399,096 pixels, or a little over 12.5%, were greater than 1. Communication with Dr. Wolter confirmed the suspected issue, the models were developed independently, leading to pixels where the balsam fir basal area model had larger amounts of basal area than the total basal area model (Wolter 2017). The balsam fir proportion raster was clipped down to the extent of Cook County, using the Extraction by Mask tool. The pixel values were then examined using the sample tool. The clipped balsam fir proportion raster had a range of 0 to 47,978.8, with 127,866, or 5.6%, of 2,294,639 pixels having a value greater than 1. The extract by attributes tool was used to remove all of the pixels with a value greater than 1. The final proportion of balsam fir raster can be seen in Figure 66 and the statistics for the raster can be viewed in Figure 67.

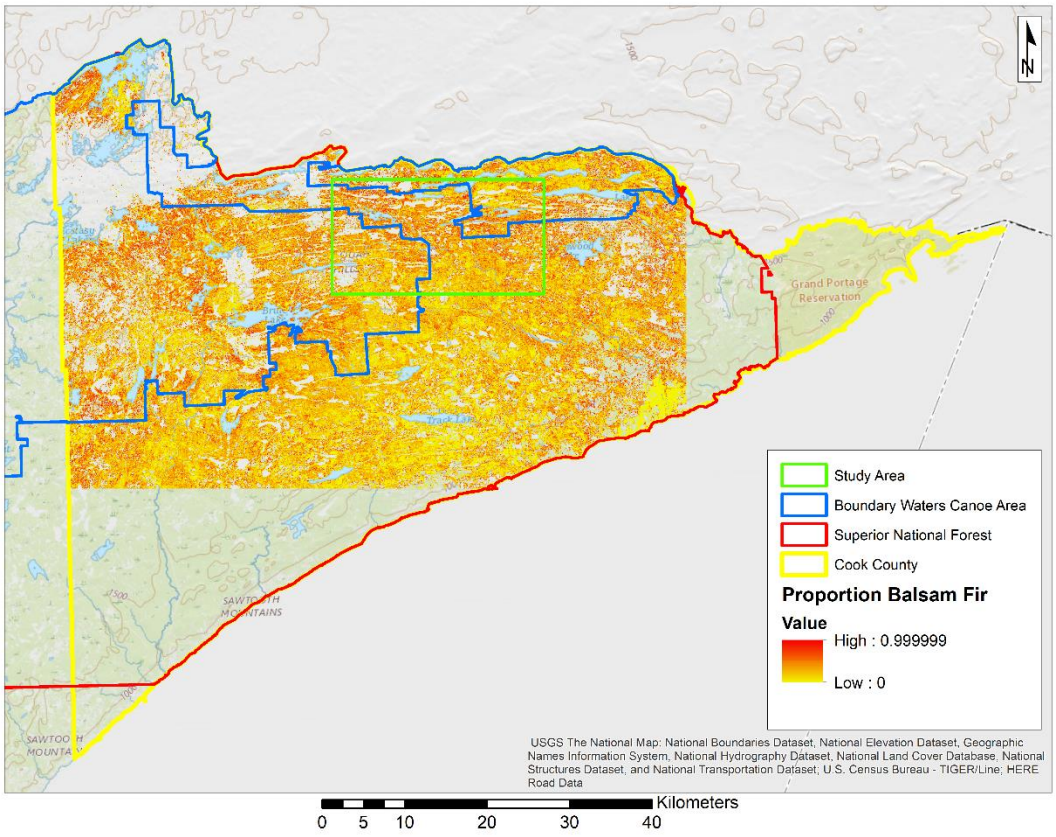

Figure 66. The final proportion of balsam fir raster after having been clipped to the extent of Cook County and having the pixels with values greater than 1 removed.

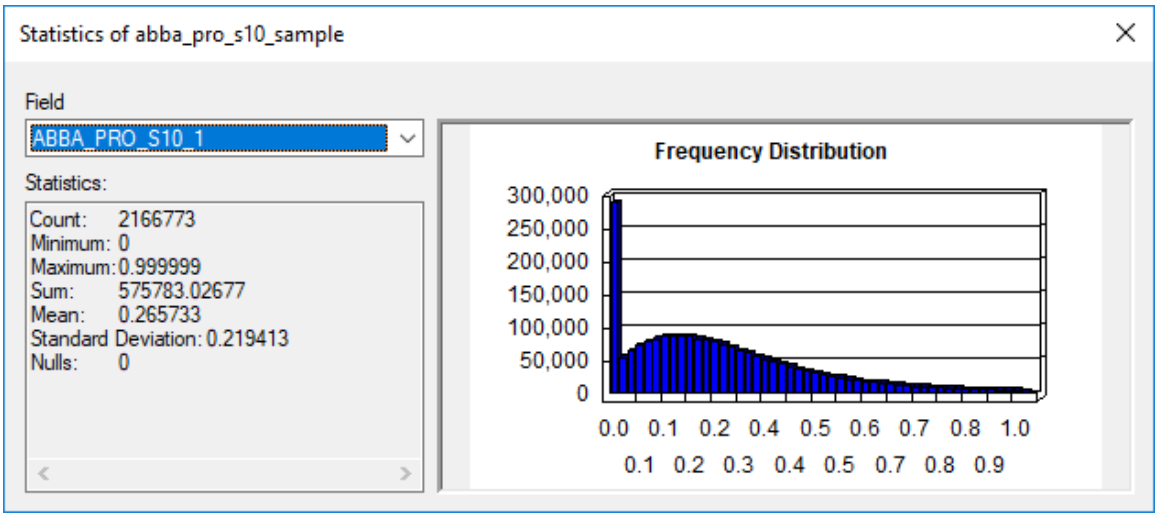

Figure 67. Statistics for the final version of the proportion of balsam fir raster.

The areas that were burned by large fires (Figures 32 and 33) also appeared to be visible in the proportion of balsam fir raster. The absence of balsam fir in those areas made sense, as the literature indicated that balsam fir is usually absent from stands for 30 to 50 years following intense fire (Uchytil 1991). The histogram also indicated that there were very few pixels approaching 100% (1 proportionally) balsam fir, this also made sense as balsam fir trees are rarely found in pure stands (Uchytil 1991).

The understory cover model (Figure 34) was multiplied by the proportion of balsam fir raster (Figure 66) to produce a balsam fir understory cover raster (Figure 68).

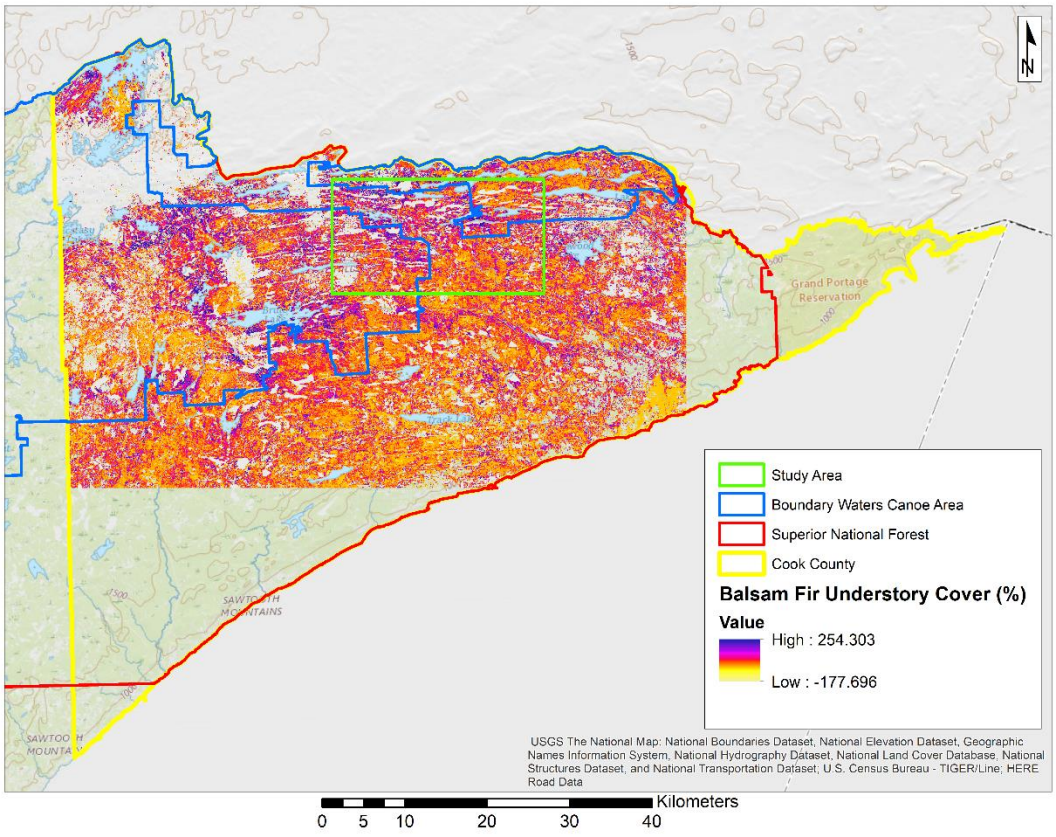

Figure 68. The raster product of the understory cover model and the proportion of balsam fir raster.

The sample tool was then used to extract the pixel values for the raster. The balsam fir understory cover raster had a range from -178% to 254%, with a mean of nearly 27% (Figure 69). The balsam fir understory cover raster had 37,971 out of 2,153,474 pixels (1.8%) with a value of less than 0% and 70,982 pixels (3.3%) greater than 100% balsam fir understory cover.

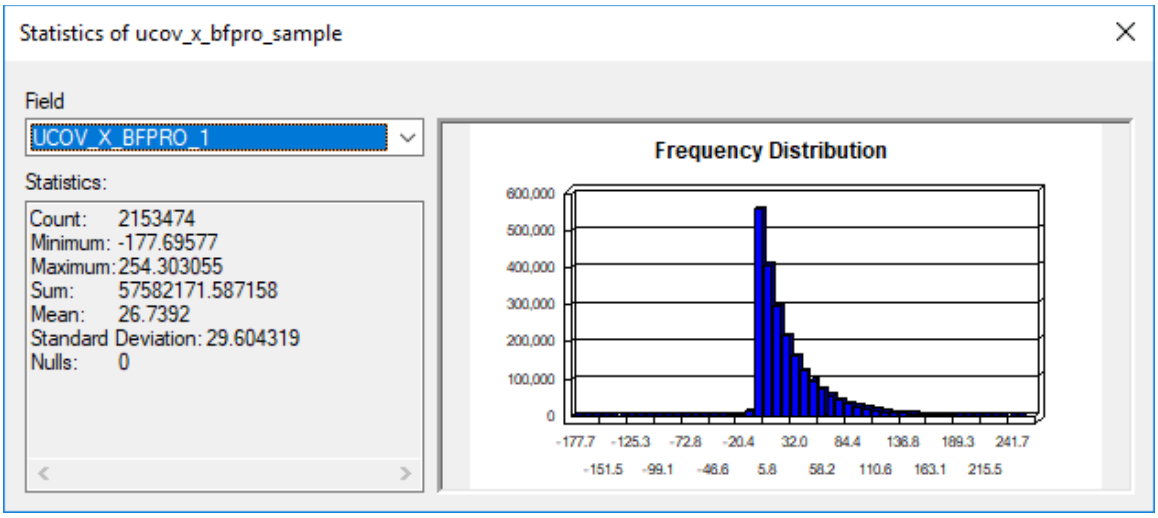

Figure 69. Statistics generated from the balsam fir understory cover raster.

For the sake of comparison, the proportion of balsam fir raster was also multiplied by the rescaled understory cover model (Figure 38). The raster product can be seen in Figure 70. The range of values for this raster was 0 – 98%, with a mean of over 16% (Figure 71). Only 79,522 of 2,153,474 pixels (3.7%) have a value of greater than 50% balsam fir understory cover. The histogram of the data appeared to be very similar in shape to that of the proportion of balsam fir raster (Figure 67).

The measured balsam fir understory cover from the May 2017 field plots was compared to both the balsam fir understory cover rasters. The predicted balsam fir understory cover values were determined from the raster pixel values at each of the plot

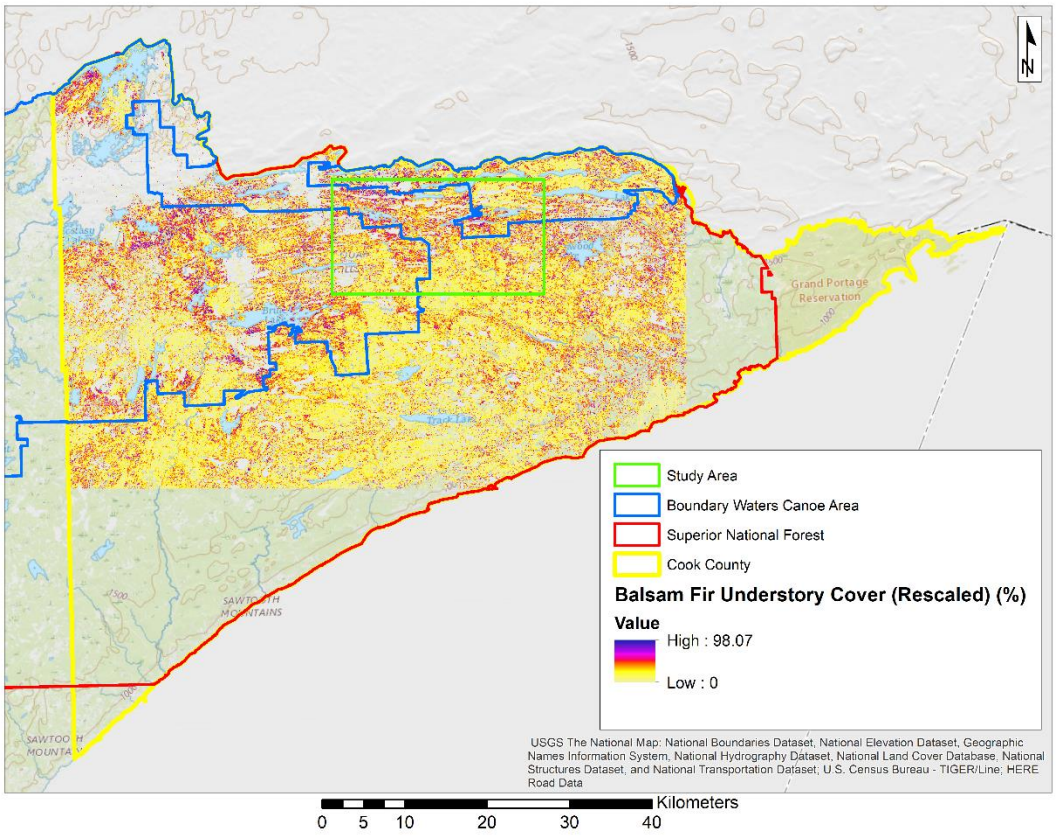

Figure 70. The raster product of the proportion of balsam fir raster and the rescaled understory cover model.

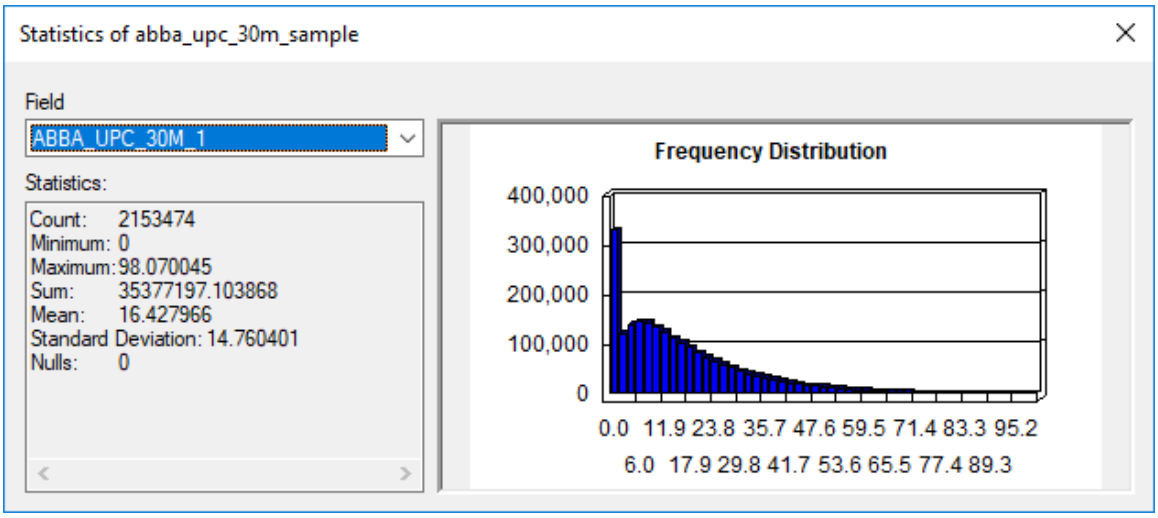

Figure 71. Statistics from the raster product of the rescaled understory cover raster and the proportion of balsam fir raster.

locations. The data is summarized in Table 11.

Plot	Measured Mean Balsam Fir Cover	Predicted Balsam Fir Understory Cover	Predicted Balsam Fir Understory Cover (Rescaled)
23	57.11%	63.98%	37.69%
30	4.92%	No Data	No Data
31	62.68%	61.36%	30.75%
42	53.63%	76.58%	35.02%
48	29.54%	9.57%	14.43%
56	43.33%	70.66%	35.32%
57	82.75%	103.33%	46.59%
59	60.52%	No Data	No Data
110	59.38%	23.47%	12.16%
139_Alt	55.98%	34.50%	16.63%
154_W40	22.80%	-0.81%	4.81%
155	48.40%	0.68%	1.48%
156	56.88%	0.54%	0.88%
166	59.20%	14.35%	13.14%
167	57.23%	26.42%	18.80%
169_Extra	10.26%	-2.78%	0.43%
170	58.70%	60.31%	41.76%
184	80.71%	17.12%	15.69%
188	86.73%	10.96%	8.53%
198	47.08%	68.95%	32.19%
199	3.64%	27.71%	12.98%
206_NW10	23.71%	24.80%	12.96%
207	32.76%	29.53%	17.60%
303	2.50%	-10.74%	2.72%

Table 11. Plot validation data for the balsam fir understory cover models.

The validation data showed that both balsam fir understory cover rasters had poor predictive power (Figures 72 and 73). The validation data for the non-rescaled balsam fir understory cover raster had a coefficient of determination of 0.1748 (correlation coefficient of 0.4181) (Figure 72, upper left) and the validation data for the rescaled balsam fir understory cover raster had a coefficient of determination of 0.1803 (correlation coefficient of 0.4246) (Figure 73, upper left).

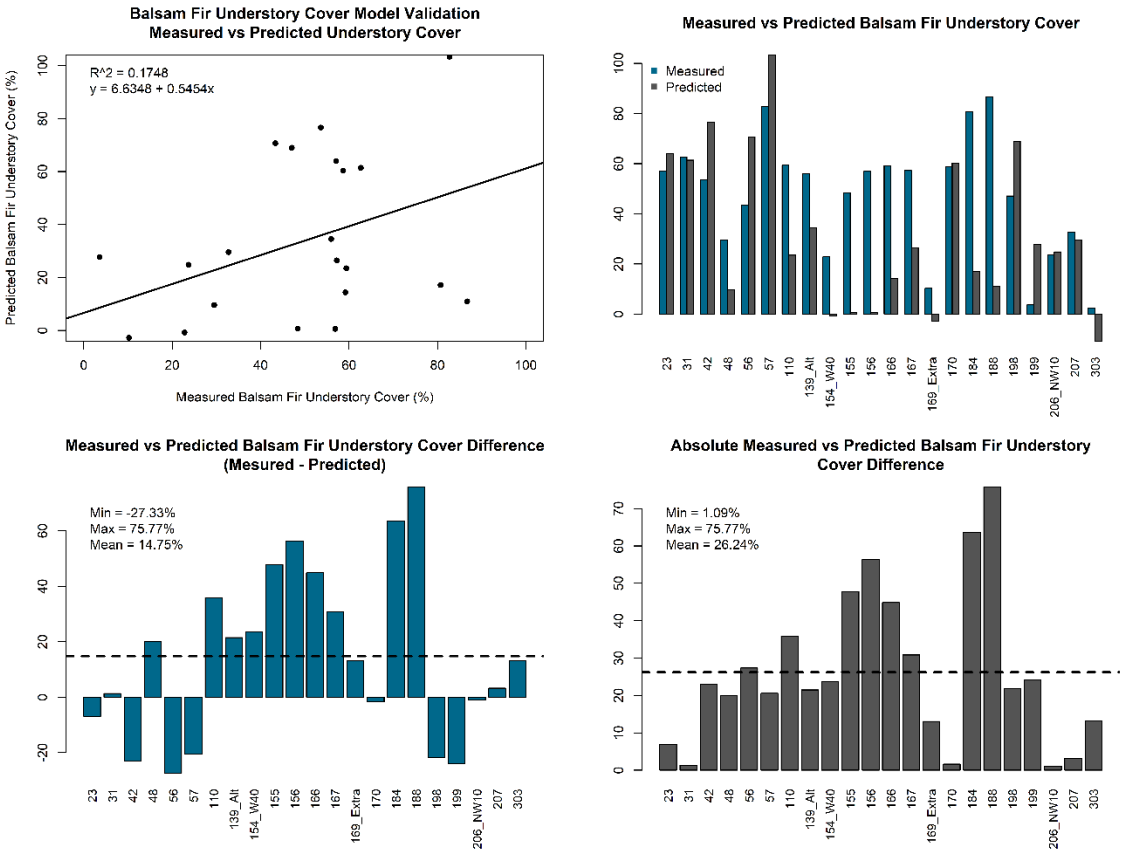

Figure 72. Graphs displaying the relationships between the measured and predicted balsam fir understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured balsam fir understory cover minus the predicted balsam fir understory cover.

The mean plot value for the difference between the measured and predicted balsam fir understory cover was 14.75% for the raster built using the non-rescaled understory cover model. This indicated that the model was under predicting the balsam fir understory cover. The upper right graph and lower left graph in Figure show that the majority of the plots had under predicted balsam fir understory cover values.

The mean plot value for the measured and predicted rescaled balsam fir understory cover difference was 28.29% (Figure 73 lower left). The rescaled balsam fir understory cover raster under predicted the balsam fir understory cover even more

consistently (Figure 73 upper right and lower left). Only two plots had over predicted balsam fir understory cover values, and it was only by 0.22% in plot 303.

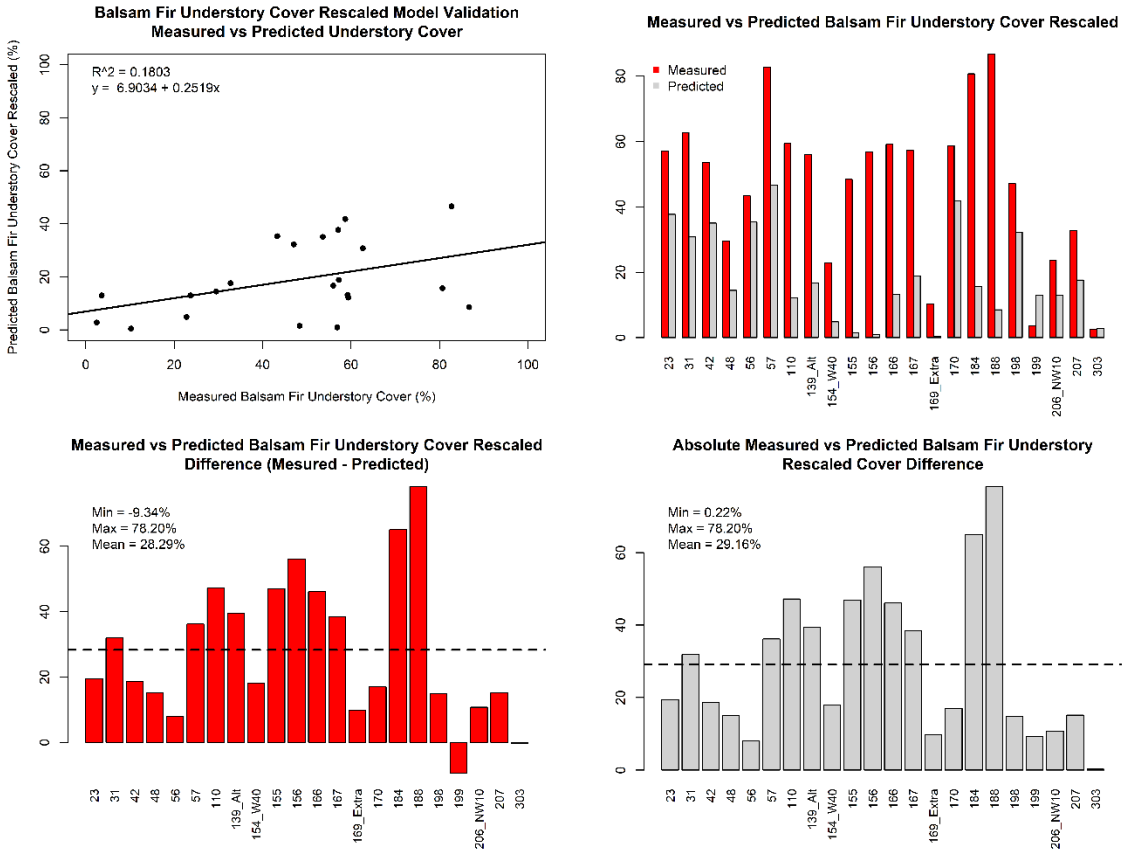

Figure 73. Graphs displaying the relationships between the measured and predicted rescaled balsam fir understory cover values at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured balsam fir understory cover minus the predicted rescaled balsam fir understory cover.

In terms of the absolute difference between the measured and predicted balsam fir understory cover, both the non-rescaled and the rescaled balsam fir understory cover rasters preformed relatively similarly. The non-rescaled balsam fir understory cover raster had a minimum absolute difference of 1.09%, compared to 0.22% for the rescaled raster. The maximum absolute difference for the non-rescaled raster was 75.77%, 2.43%

less than the maximum absolute difference of 78.20% for the rescaled balsam fir understory cover raster. The mean absolute difference was also similar for both rasters. The mean absolute difference was 26.24% for the non-rescaled raster and 29.16% for the rescaled raster.

The poor predictive power of both balsam fir understory cover models was possibly because of the understory cover model and the basal area models (Wolter 2016b) measuring two different things in two different ways. The cover line intercept data that went into building the understory cover model measured overlapping canopies from trees that were of the same species as one continuous canopy, whereas the basal area models' (Wolter 2016b) field data included every measurable tree in each plot in the model. In many of the plots, balsam fir was found in clusters, and each tree's basal area would have been included if measured using the methods of Wolter (2016b), but not with the cover line intercept method.

For example, in plot 139_Alt, which had a mean measured balsam fir understory cover of 55.98%, transect 2 had 3 intercepts with balsam fir. The first intercept went from 1.31 to 4.48 meters and was comprised of 4 saplings and 4 seedlings. The second balsam fir intercept started at 6.68 meters, was 1.74 meters in length, and included 1 sapling and 2 seedlings. The final balsam fir intercept in plot 139_Alt, transect 2 crossed 14 saplings and 11 seedling crowns, started at 8.93 meters, and went to the end of the transect.

Another issue could be that because balsam fir is typically found in the understory; using the direct proportion of balsam fir with the understory cover models might not be the correct approach. In both of the balsam fir understory cover rasters, the

non-rescaled and the rescaled, the amount of balsam fir understory cover was under predicted in most plots. A correction factor might need to be considered in future work.

Summary and Conclusion

In order to explore the possibility of modeling the amount of balsam fir in the understory of forest stands in Superior National Forest, a collaboration with Dr. Peter Wolter of Iowa State University was established. Dr. Wolter mapped the basal area of several tree species in the study area of this investigation (Wolter 2016b). The balsam fir basal area model (Wolter 2016b) was compared with the understory cover model through binary rasters. Results of that comparison indicated a possible correlation between high amounts of balsam fir basal area and moderate to high amounts of understory cover.

An attempt was also made to combine the understory cover model with the balsam fir basal area model. Both the original and rescaled versions of the understory cover model were considered. Validation of the results with the May 2017 cover line intercept data showed that both versions had relatively poor predictive power. The correlation coefficient for the balsam fir understory cover model was 0.4181 and the correlation coefficient for the rescaled balsam fir understory cover model was 0.4246. Plot level comparisons indicate that both models were under predicting balsam fir cover at most plots.

ADDITIONAL DEVELOPMENT OF UNDERSTORY COVER MODELS

Introduction

The results from the validation of the original understory cover model were not as good as anticipated. Although not tested, one possible reason might have been a statistical difference between the plots that the model was built from and the plots that the model was validated with. The mean plot understory cover percentages ranged from 12.42% to 199.52%, with a mean of 69.40% for the understory cover model's training plots (Table 5). The mean plot understory cover values for the testing (validation) plots ranged from 35.98% to 158.01%, with a mean value of 98.10% (Table 6).

The initial thought was that perhaps the plots in which there was only one transect measured had a major influence on the model. The original understory cover model was built from 45 plots, in which a total of 114 transects had been measured, meaning that the model was built from 23 plots had been measured with four transects and 22 plots that had only one transect measured. The model was then validated with data from 23 plots that had 4 measured transects and 1 plot that had only 1 transect measured.

Another thought was that perhaps that since very few plots with either very high or very low understory covers were measured in the field, out of all of the plots, those high and low plots might not be representative of the actual conditions in the study area and might be affecting the results.

To investigate these issues, a series of models were constructed from various selections of training and testing plots from the pooled data. These selections were used to determine if a better model could be constructed from, what might be, a more representative set of the data. Several strategies were tested, including constructing a

model from the four transect plots and comparing it to the one transect plots. The remainder of this chapter outlines the strategies and the results of each modelling attempt.

Materials and Methods

Basic Statistical Information and Data Stratification

To help facilitate the model building, the plot data (Tables 5 and 6) were pooled together. From that pooled data some basic statistics were calculated, the mean and standard deviation. A histogram, a boxplot, and a number line of the data were also generated. The mean and the standard deviation were used to divide the plots into three groups. This data stratification is based on the work of Wylie et al. (2008), Gu and Wylie (2015), Boyte, Wylie, and Major (2015), and Gu and Wylie (2017). The goal of the data stratification was to be able to build models with a better representation of existing conditions.

Stratified Random Sample Model

In the first strategy employed for additional model building, two-thirds of the plots from each level of stratification were randomly selected to be used as training plots, leaving one-third from each level to be used as test plots. Five test plots were randomly drawn from the low level, 15 test plots were randomly drawn from the medium level, and 3 test plots were randomly drawn from the high level. This left 9 low level plots, 31 medium level plots, and 7 high level plots for training data. From the stratified plots, three sets of randomly selected training and test plots were drawn. Each of the three randomly selected groups of training plots were compared to the plot level lidar metrics

using multiple linear regression. The model with the highest coefficient of determination was then output spatially. The model's predicted pixel values were then compared to the test plot data for validation.

Stratified 75% Medium Level Plots and 35% High and Low-Level Plots Model

A second modeling strategy was employed, again based on the work of Wylie et al. (2008). In this strategy, a model was developed by using 75 percent of the medium level plots, 35 percent of the low-level plots, and 35 percent of the high-level plots as training plots. This equated to five low level, thirty-five medium level, and four high level plots for model building and nine low level, eleven medium level, and six high level plots for validation. Again, three random samples were drawn. These random samples were compared to the lidar level lidar metrics using multiple regression analysis. The model with the highest coefficient of determination was rasterized and the pixel values at the validation plot locations were determined. These predicted values were then compared to the field measured plot values.

80% Training Data and 20% Test Data Model

Gu et al. (2016) determined that using 80% training data and 20% test data produced "optimal" results when building regression tree models with remotely-sensed data. The next employed strategy was based on that idea. Three randomly selected groups of 80% training plots and 20% test plots (56 training and 14 test plots) were generated. The training plots were compared to the plot level lidar metrics via multiple regression analysis. A raster was built from the raster with the highest coefficient of

determination. The predicted model values were determined at each of the test plot locations from the raster. The predicted model values were compared to the test values.

80% Training Data and 20% Test Data Model Withholding Plot 62 from Training

An additional model was built from the 80% training data and 20% test data modeling approach. From the second set of randomly generated test plots, a randomly selected plot was chosen to replace plot 62 in the training set. The new set of training plots was then used to generate a model through multiple regression and the model equation was used to generate a spatial output. From that spatial output, the predicted plot values were determined for the test plots. The predicted values were then compared to the field measured values of the test plots.

70% Training Data and 30% Test Data Model

For the sake of comparison, a model was built with 70% training plots and 30% test plots. As before, the three randomly selected groups of training and test plots were generated. Each group contained 49 training plots and 21 test plots. Multiple regression analysis was used to compare the field measurements of the training plots to the lidar metrics for each plot. The group of training plots that produced the highest coefficient of determination was then used to generate a raster. From the raster, the predicted understory cover values for the test plot locations were determined. Those predicted values were then compared to the measured understory cover values for each test plot.

4 Transect Plots for Training and 1 Transect Plots for Testing Model

It was thought, that perhaps, the plots in which only 1 transect was measured might not be representing as much of the variability in each plot as the plots in which 4 transects were measured. So, a model was built using the 46) 4-transect plots as training data and the 24) 1-transect plots as test data. The training plot understory cover measurements were compared to the plot level lidar metrics via multiple regression analysis. The equation of the resulting model was used to generate a spatial version of the model. From the spatial model, the predicted understory cover values of the test plots were determined. The measured and predicted values of understory cover were then compared for validation of the model.

Original Model with Plot 121

During the generation of the original understory cover model, plot 121 was left out of the training data set because plot 121 did not have a phase 1 measurement. The original understory cover model was rebuilt, using multiple regression, with plot 121 in the training data. The revised model was output spatially and the same test data set as the original model was used for validation.

Statistics Generated from the Rasterized Models

Following the building of all eight models (the seven outlined here plus the original), the cell statistics tool was used to generate several raster layers with information concerning the eight models. The cell statistics tool, "calculates a per-cell statistic from multiple rasters," (Esri Inc. 1999-2017) and outputs a raster with those

statistical values. For all eight of the models, the majority (which is the mode), the maximum, the mean, the median, the minimum, the range, and the standard deviation were calculated. The raster calculator tool was then used to compare the raster of each model to the mean and median rasters. This was done by subtracting the mean and median rasters from each of the model rasters. The median raster was also subtracted from the mean raster. The extraction by mask tool was used to clip each raster that was produced in this exercise down to the extent of Cook County and to remove the water areas from each raster.

Results and Discussion

Basic Statistical Information and Data Stratification

All 70-understory cover (phase 2) plots were pooled together. The overall range in understory cover for all 70 plots was 12.42% to 199.52%, with a mean of 79.73%, and a standard deviation of 42.52%. To help visualize the data, a histogram (Figure 74), boxplot (Figure 75), and number line (Figure 76) were generated.

The histogram (Figure 74) made it appear as though there were few plots with understory cover less than 20% and few plots with understory cover greater than 160%. There were also a couple of dips in the data between 60 and 80% understory cover and between 120 and 140% understory cover. The boxplot (Figure 75) confirmed that many of the data points fell between 40 and 120% understory cover. Gaps could also be seen in the number line (Figure 76) between 0 and 10%, 18 and 26%, 70 and 80%, 130 and 140%, 145 and 155%, 168% and 175%, and 180 and 199%. Plot 62 appeared to be an outlier, at 199.52% understory cover (Figure 75, Table 5).

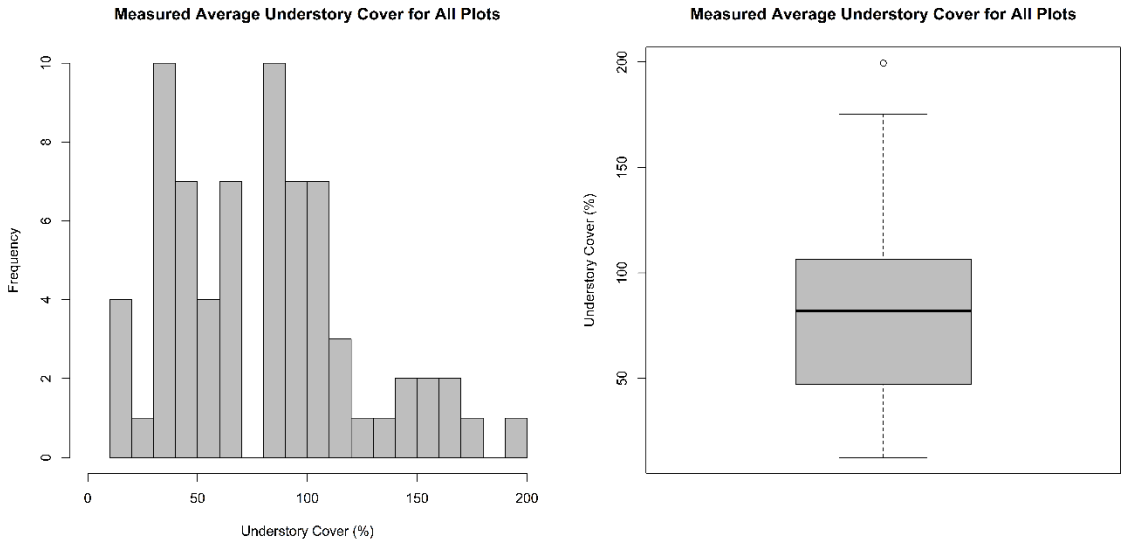

Figure 74 (Left) and Figure 75 (Right). Figure 74 displays a histogram of the measured plot understory cover for all 70 plots. Figure 75 displays a boxplot of the same data.

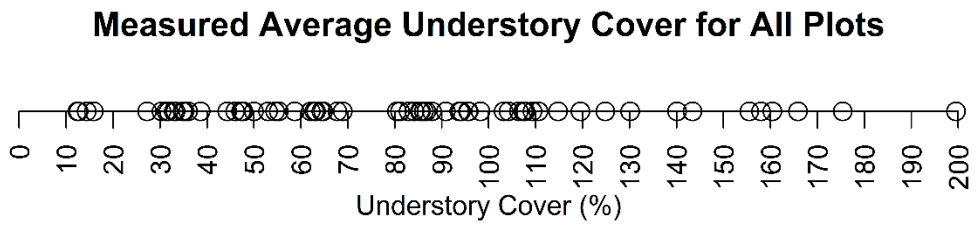

Figure 76. A number line showing the mean plot understory for each of the 70 plots.

The 70 plots were then split into three groups. Plots that were more than one standard deviation less than the mean (less than 37.21% understory cover) were placed into the low category. Plots that were more than one standard deviation greater than the mean (greater than 122.26% understory cover) were placed into the high category. The remainder of the plots, those within one standard deviation of the mean, composed the medium category. There were fourteen plots that made up the low category, forty-six plots made up the medium category, and ten plots in the high category (Table 12).

Low Understory Cover	Medium Understory Cover	High Understory Cover
90, 95, 98, 99, 106, 108, 141, 150, 151, 169_Extra, 211, 229, 230, 242	23, 27, 30, 31, 56, 57, 59, 79, 80, 84, 89, 103, 104, 105, 107, 121, 122, 139_Alt, 142, 143, 152, 153, 154_W40, 155, 156, 166, 167, 168, 184, 188, 198, 199, 206, 207, 210, 212, 214, 222, 223, 228, 239, 240, 241, 243, 247, 303	13, 14, 16, 26, 42, 48, 62, 92, 110, 170

Table 12. Plots in each division of understory cover.

Stratified Random Sample Model

In the stratified random sampling effort, the first random sample generated the highest coefficient of determination, 0.5551, in the multiple regression analysis. The equation generated by the model was:

$$UC = 49.2972 + -9.2866*p50 + -3.9628*d8.10m + 1.9076*TotCov.$$

where UC was the understory cover, p50 was the height of the 50th percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. This equation was used to output the model spatially. The pixel values were determined at the location of each of the randomly chosen test plots. Those predicted understory cover values were then compared to the measured understory cover values for validation (Figure 77).

The coefficient of determination for the measured versus predicted understory cover was 0.3567 (correlation coefficient of 0.5972) (Figure 77, upper left). The minimum difference between the measured and predicted understory cover was -48.80%, with a maximum difference of 52.91%, and a mean difference of -0.43% (Figure 77, lower left). The minimum absolute difference between the measured and predicted values was 1.61%, with a maximum absolute difference of 52.91%, and a mean absolute difference of 28.89% (Figure 77, lower right).

Figure 77. Graphs displaying the relationships between the measured and predicted understory cover values for the stratified random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

Stratified 75% Medium Level Plots and 35% High and Low-Level Plots Model

In the stratified 35% low level plots, 75% medium level plots, and 35% high level plots modeling strategy, the second random sample produced the model with the highest coefficient of determination. That coefficient of determination was 0.4285 and the equation for the model was:

$$UC = 57.740 + -12.930*p40 + 5.491*d14.16m + 1.170*TotCov$$

where UC was the understory cover, p40 was the height of the 40th percentile of returns, d14.16m was the relative density between the heights of 14 and 16 meters, and TotCov was the total cover metric. The equation was used to generate a spatial output of the model. The model's predicted values were determined at the locations of the test plots and comparison with the measured values gave a coefficient of determination of 0.2547 (correlation coefficient of 0.5047) (Figure 78, upper left). The range of values between

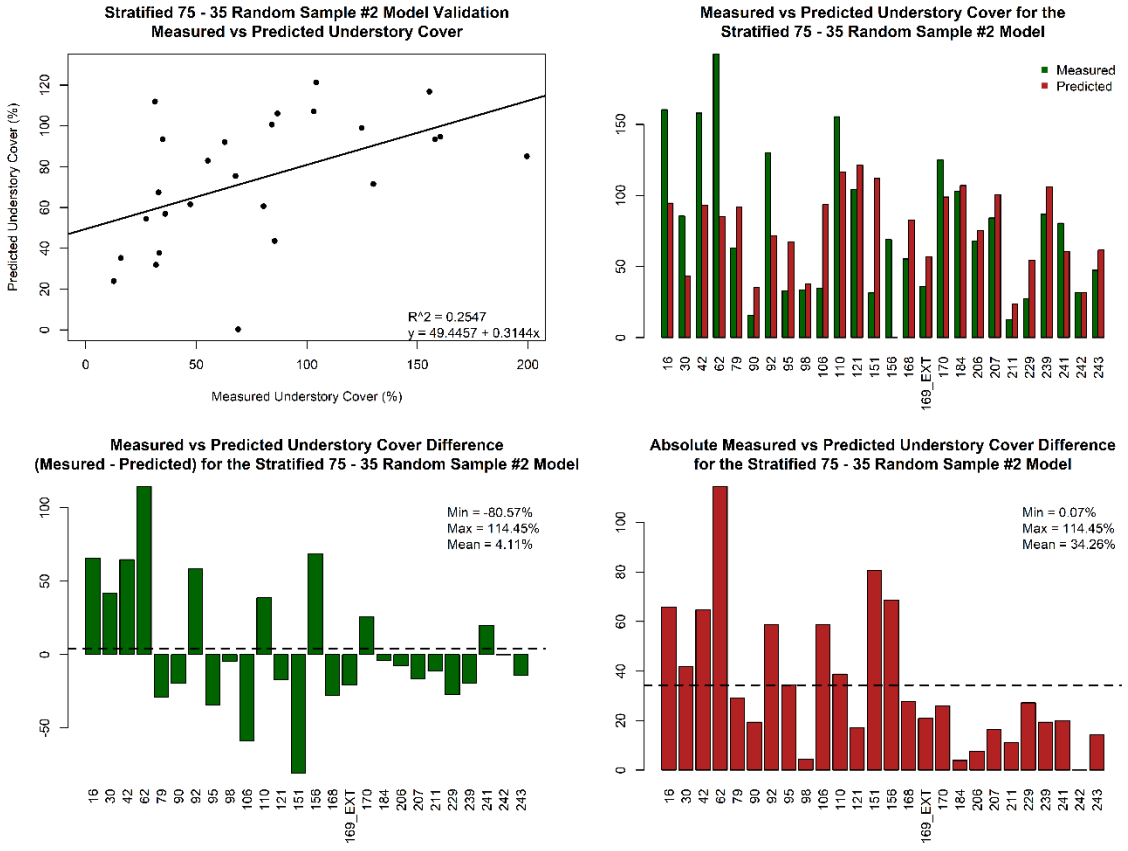

Figure 78. Graphs displaying the relationships between the measured and predicted understory cover values for the stratified 75% of medium and 35% of high and low plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

the measured and predicted amounts of understory cover was -80.57% to 114.45%, with a mean of 4.11% (Figure 78, lower left). The absolute difference between the measured and predicted understory cover ranged from 0.07% to 114.45%, with a mean absolute difference of 34.26% (Figure 78, lower right).

80% Training Data and 20% Test Data Model

The third random sample produced the highest coefficient of determination (0.5035) (Figure 79, upper left) in the 80% training data and 20% test data modeling

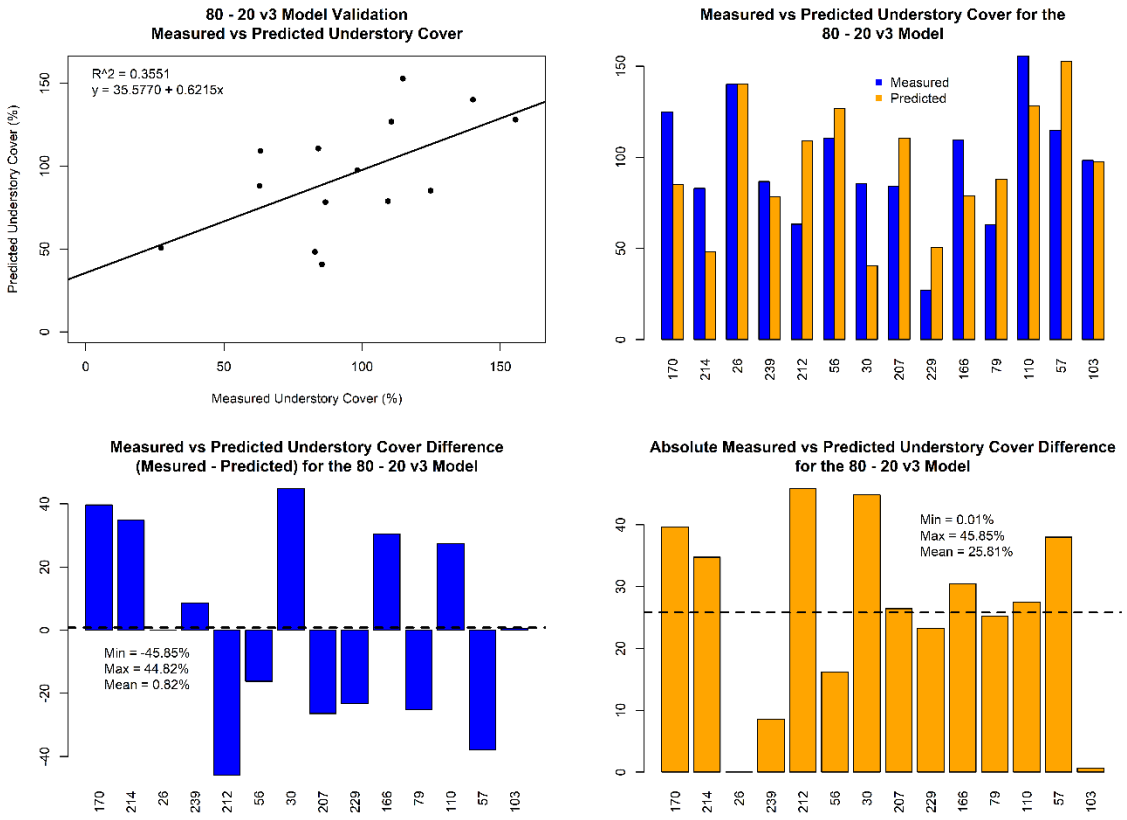

Figure 79. Graphs displaying the relationships between the measured and predicted understory cover values for the 80% training and 20% test plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

approach. The resulting equation from the multiple regression modeling was:

UC = 37.5703 + -8.1154*p50 + -4.2110*d8.10m + 1.9709*TotCov.

where UC was the understory cover, p50 was the height of the 50th percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. When the predicted and measured test plot understory covers were compared, the results produced a coefficient of determination of 0.3551 (correlation coefficient of 0.5959) (Figure 79 upper left). The measured and predicted understory covers had a minimum difference of -45.85% and a maximum difference of 44.82%, with a mean difference of 0.82% (Figure 79, lower left). The minimum absolute difference between the measured and predicted values was 0.01%, the maximum was 45.85%, and the mean was 25.81% for the validation plots (Figure 79, lower right).

80% Training Data and 20% Test Data Model Withholding Plot 62 from Training

Plot 62 was chosen to be removed from the training data as it appears to be an outlier in the boxplot of the measured plot understory covers (Figure 75). The equation produced by this modeling attempt was:

UC = 41.6575 + -7.3908*p50 + -3.3324*d8.10m + 1.7087*TotCov

where UC was the understory cover, p50 was the height of the 50th percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. This version of the 80% training data and 20% test data model had a coefficient of determination of 0.5223, which was indeed higher than the 80% training data and 20% test data model (0.5035) that contained plot 62 in the training data set.

When comparing the measured and predicted understory covers for the test plots, a coefficient of determination of 0.3721 (correlation coefficient of 0.6100) resulted (Figure 80, upper left). The minimum difference between the measured and predicted understory cover was -57.33%, the maximum was 85.54%, and the mean difference was 8.38% (Figure 80, lower left). The minimum absolute difference between the measured and predicted understory cover was 3.45%, the maximum absolute difference was 85.54%,

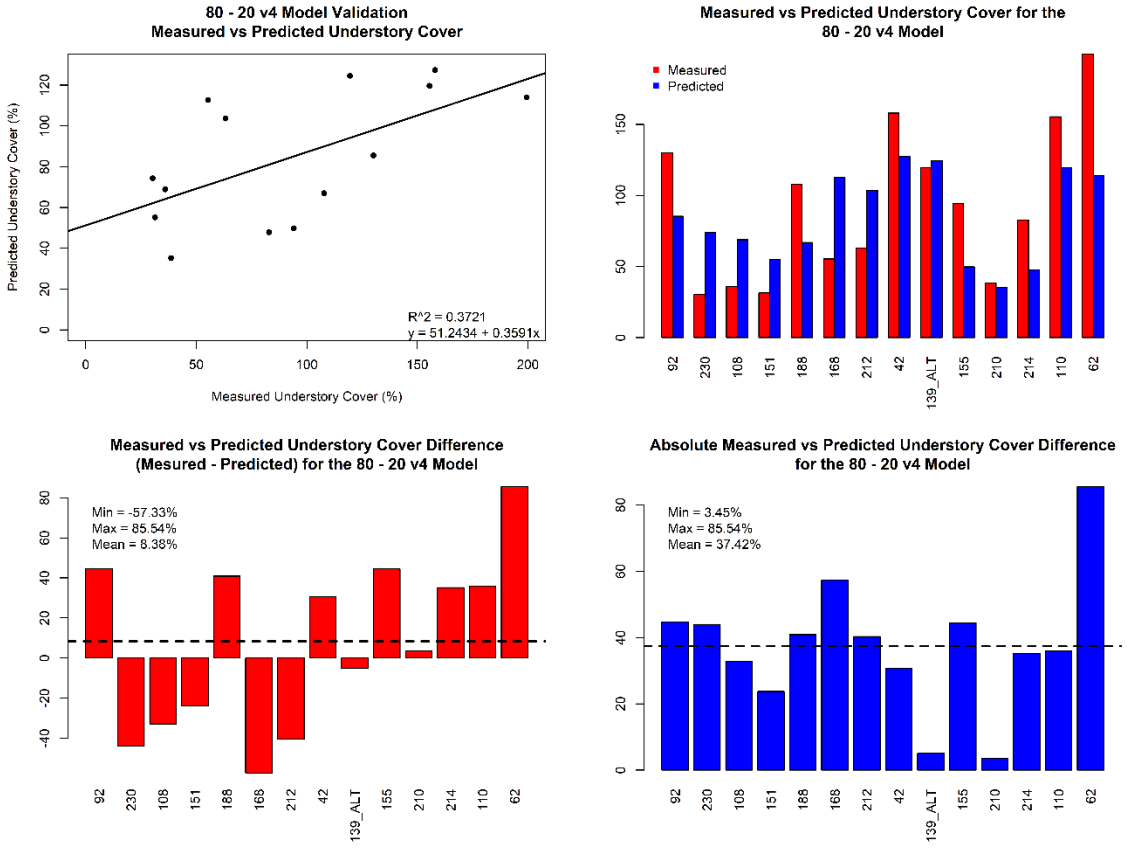

Figure 80. Graphs displaying the relationships between the measured and predicted understory cover values for the 80% training and 20% test plots, withholding plot 62 from the training data, random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

and the mean absolute difference was 37.42% (Figure 80, lower right). While both the coefficient of determination for the model and the coefficient of determination for the validation data were higher for this model than for the 80% training data and 20% containing plot 62 in the training data set, this version of the model had both a higher mean difference and mean absolute difference when comparing the measured and predicted understory cover for the test plots.

70% Training Data and 30% Test Data Model

In the case of the 70% training data and 30% test data model, the third random sample generated the highest coefficient of determination, 0.5133. The resulting model had an equation of:

$$UC = 48.7549 + -8.3475*p50 + -3.5195*d8.10m + 1.8221*TotCov$$

where UC was the understory cover, p50 was the height of the 50^{th} percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. When the measured and predicted values were plotted, the results produced a coefficient of determination of 0.4739 (correlation coefficient of 0.6884) (Figure 81, upper left). The minimum difference in measured and predicted understory cover values was -69.44%, with a maximum of 42.60%, and a mean difference of -19.44% (Figure 81, lower left). When compared, the measured and predicted understory cover values for the test plots had absolute differences that ranged from 3.28% to 69.44%, with a mean absolute difference of 26.06% (Figure 81, lower right).

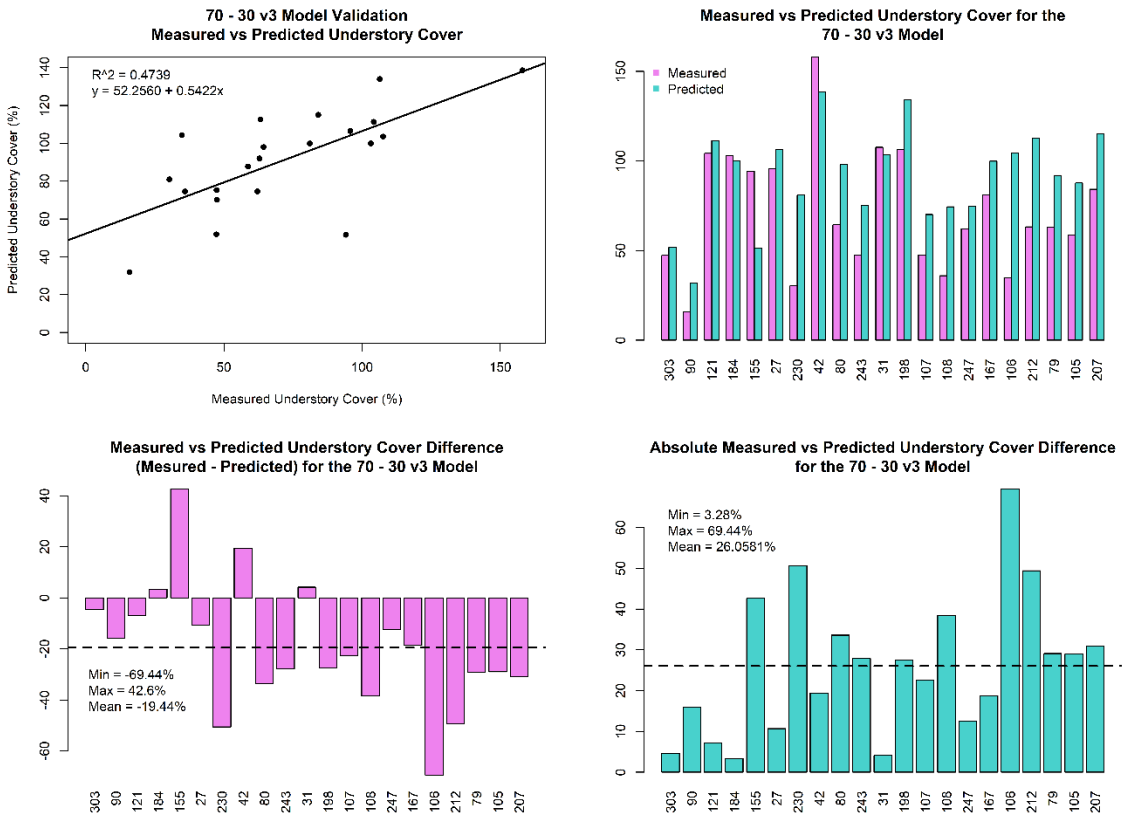

Figure 81. Graphs displaying the relationships between the measured and predicted understory cover values for the 70% training and 30% test plots random sample model at each validation plot. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

4 Transect Plots for Training and 1 Transect Plots for Testing Model

The model built from the 4 transect plots produced a coefficient of determination of 0.4619. The model's equation was:

$$UC = 40.7542 + -4.7480*p50 + -2.9639*d8.10m + 1.3574*TotCov$$

where UC was the understory cover, p50 was the height of the 50th percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. Comparing the measured and predicted understory cover

percentages for the 1 transect test plots produced a coefficient of determination of 0.5828 (correlation coefficient of 0.7634) (Figure 82, upper left). A minimum difference of -51.66%, a maximum difference of 97.09%, and a mean difference of 8.79% were calculated by comparing the measured and predicted understory cover values from the test plots (Figure 82, lower left). The absolute differences ranged from 0.33% to 97.09% and the mean absolute difference was 33.47% for the measured and predicted understory cover values in the test plots (Figure 82, lower right).

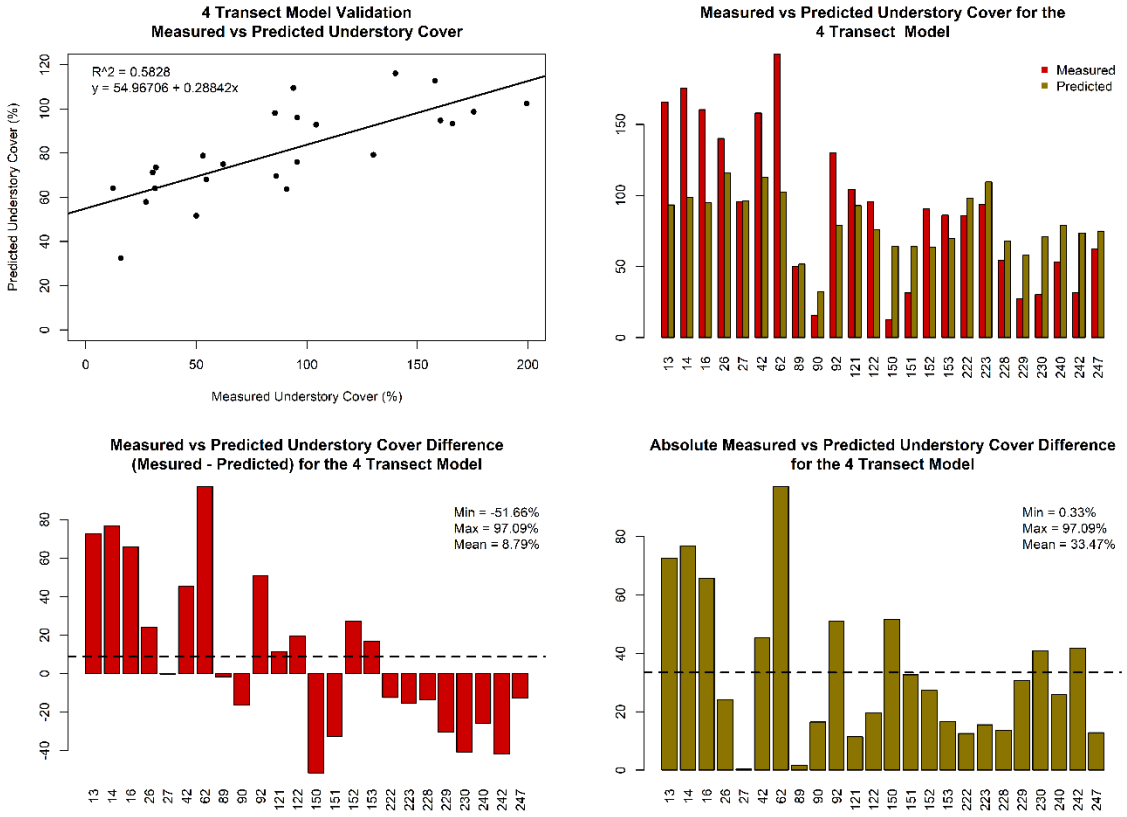

Figure 82. Graphs displaying the relationships between the measured and predicted understory cover values for the validation data for the model built from the 4 transect plots. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

Original Model with Plot 121

Adding plot 121 into the training data set for the original model produced a revised model with a coefficient of determination of 0.6547. This was higher than the coefficient of determination of the original model, which was 0.6504. The equation for the revised model appeared to be very similar to that of the original model and was:

UC = 36.5579 + -12.3116*p50 + -9.0397*d8.10m + 2.8275*TotCov

where UC was the understory cover, p50 was the height of the 50th percentile of returns, d8.10m was the relative density between the heights of 8 and 10 meters, and TotCov was the total cover metric. When the measured and predicted amounts of understory cover for the test plots were plotted against each other, a coefficient of determination of 0.4698 (correlation coefficient of 0.6854) resulted (Figure 83, upper left). This was the same as the original model. The range of differences between the measured and predicted understory cover values for the test plots was -83.55% to 122.67%, with a mean difference of 14.18% (Figure 83, lower left). The range was narrower than the original model and the mean difference was slightly less than the original, which was 14.30%. The minimum mean absolute difference between the measured and predicted understory cover percentages was 0.22%, the maximum absolute difference was 122.67%, and the mean absolute difference was 46.88% (Figure 83, lower right). Again, these values were slightly less than that of the original, which had a minimum of 0.38%, a maximum of 123.09%, and a mean of 47.04%.

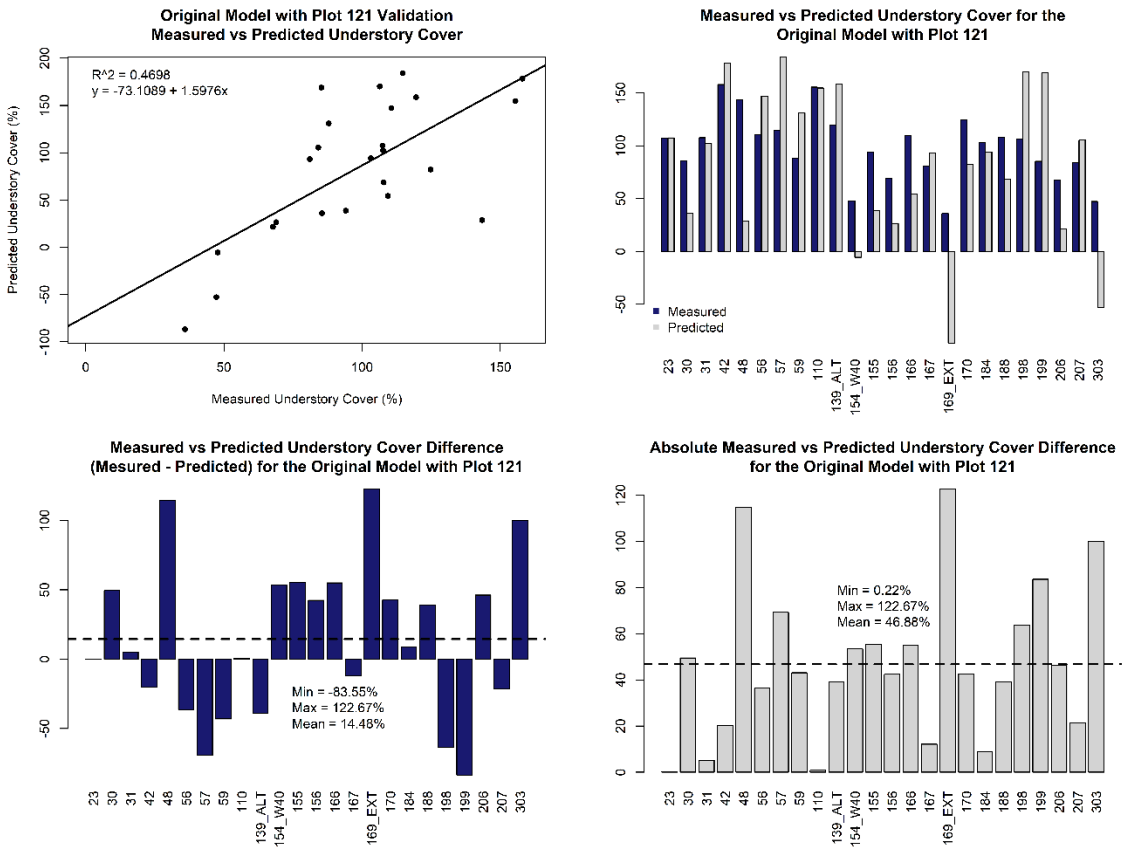

Figure 83. Graphs displaying the relationships between the measured and predicted understory cover values at each validation plot for the revised original model containing plot 121 in the training data. The upper left graph displays the plot of measured vs predicted values. The upper right graph shows a side by side comparison of the measured versus predicted values. The lower right graph shows the absolute difference between the measured and predicted values. The lower left graph shows the difference between the measured and predicted values where the difference is determined by the measured understory minus the predicted understory cover.

Comparison of the Models

A table was created (Table 13) in order to facilitate the comparison of all eight models (the seven outlined above, plus the original model). The models had coefficients of determination that ranged from 0.4285 to 0.6547, with the original model with plot 121 added to the training data having the highest coefficient of determination and the original model itself having the second highest. This was an interesting result, it

Model	Original	Stratified Random Sample
Training Plots	13, 14, 16, 26, 27, 62, 79, 80, 84, 89, 90, 92, 95, 98, 99, 103, 104, 105, 106, 107, 108, 122, 141, 142, 143, 150, 151, 152, 153, 168, 210, 211, 212, 214, 222, 223, 228, 229, 230, 239, 240, 241, 242, 243, 247	14, 26, 30, 31, 42, 48, 56, 62, 80, 84, 92, 98, 99, 103, 104, 105, 106, 107, 108, 110, 121, 122, 139_Alt, 143, 150, 151, 153, 154_W40, 155, 166, 167, 168, 169_Extra, 184, 199, 206, 207, 210, 211, 222, 223, 228, 229, 239, 240, 243, 303
Proportion 1 Transect Training Plots	1 Trans = 22, 4 Trans = 23, Total = 45, Proportion of 1 Trans Plots = 0.489	1 Trans = 15, 4 Trans = 32, Total = 47, Proportion of 1 Trans Plots = 0.319
R^2	0.6504	0.5551
Variables	p50, d8.10m, TotCov	p50, d8.10m, TotCov
Equation	UC = 36.5741 + -12.3610*p50 + -9.0716*d8.10m + 2.8380*TotCov	UC = 49.2972 + -9.2866*p50 + -3.9628*d8.10m + 1.9076*TotCov
Test Plots	23, 30, 31, 42, 48, 56, 57, 59, 110, 139_Alt, 154_W40, 155, 156, 166, 167, 169_Extra, 170, 184, 188, 198, 199, 206, 207, 303	242, 90, 95, 230, 141, 152, 27, 241, 59, 212, 142, 188, 214, 247, 89, 198, 156, 79, 23, 57, 170, 13, 16
Proportion 1 Transect Test Plots	1 Trans = 1, 4 Trans = 23, Total = 24, Proportion of 1 Trans Plots = 0.042	1 Trans = 9, 4 Trans = 14, Total = 23, Proportion of 1 Trans Plots = 0.391
Validation R^2	0.4698	0.3567
Validation Equation	y = -73.5142 + 1.6036x	y = 41.1845 + 0.4995x
Min Diff M - P UC	-84.04%	-48.80%
Max Diff M - P UC	123.09%	52.91%
Mean Diff M - P UC	14.30%	-0.43%
Min Abs Diff M - P UC	0.38%	1.61%
Max Abs Diff M - P UC	123.09%	52.91%
Mean Abs Diff M - P UC	47.04%	28.89%

Table 13. Pertinent information about each of the eight models built in this investigation. One transect plots are colored in red.

Model	Stratified 75% Medium and 35% Low and High Training Plots	80% Training Plots and 20% Test Plots
Training Plots	13, 14, 23, 26, 27, 31, 48, 56, 57, 59, 80, 84, 89, 99, 103, 104, 105, 107, 108, 122, 139_Alt, 141, 142, 143, 150, 152, 153, 154_W40, 155, 166, 167, 188, 198, 199, 210, 212, 214, 222, 223, 228, 230, 240, 247, 303	13, 14, 16, 23, 27, 31, 42, 48, 59, 62, 80, 84, 89, 90, 92, 95, 98, 99, 104, 105, 106, 107, 108, 121, 122, 139_Alt, 141, 142, 143, 150, 151, 152, 153, 154_W40, 155, 156, 167, 168, 169_Extra, 184, 188, 198, 199, 206, 210, 211, 222, 223, 228, 230, 240, 241, 242, 243, 247, 303
Proportion 1 Transect Training Plots	1 Trans = 15, 4 Trans = 29, Total = 44, Proportion of 1 Trans Plots = 0.341	1 Trans = 22, 4 Trans = 34, Total = 56, Proportion of 1 Trans Plots = 0.393
R^2	0.4285	0.5035
Variables	p40, d14.16m, TotCov	p50, d8.10m, TotCov
Equation	UC = 57.740 + -12.930*p40 + 5.491*d14.16m + 1.170*TotCov	UC = 37.5703 + -8.1154*p50 + -4.2110*d8.10m + 1.9709*TotCov
Test Plots	16, 30, 42, 62, 79, 90, 92, 95, 98, 106, 110, 121, 151, 156, 168, 169_Extra, 170, 184, 206, 207, 211, 229, 239, 241, 242, 243	170, 214, 26, 239, 212, 56, 30, 207, 229, 166, 79, 110, 57, 103
Proportion 1 Transect Test Plots	1 Trans = 9, 4 Trans = 17, Total = 26, Proportion of 1 Trans Plots = 0.346	1 Trans = 2, 4 Trans = 12, Total = 14, Proportion of 1 Trans Plots = 0.143
Validation R^2	0.2547	0.3551
Validation Equation	y = 49.4457 + 0.3144x	y = 35.5770 + 0.6215x
Min Diff M - P UC	-80.57%	-45.85%
Max Diff M - P UC	114.45%	44.82%
Mean Diff M - P UC	4.11%	0.82%
Min Abs Diff M - P UC	0.07%	0.01%
Max Abs Diff M - P UC	114.45%	45.85%
Mean Abs Diff M - P UC	34.26%	25.81%

Table 13. Continued.

Model	80% Training Plots (without plot 62) and 20% Test Plots	70% Training Plots and 30% Test Plots
Training Plots	13, 14, 16, 23, 26, 27, 30, 31, 48, 56, 57, 59, 79, 80, 84, 89, 90, 95, 98, 99, 103, 104, 105, 106, 107, 121, 122, 141, 142, 143, 150, 152, 153, 154_W40, 156, 166, 167, 169_Extra, 170, 184, 198, 199, 206, 207, 211, 222, 223, 228, 229, 239, 240, 241, 242, 243, 247, 303	13, 14, 16, 23, 26, 30, 48, 56, 57, 59, 62, 84, 89, 92, 95, 98, 99, 103, 104, 110, 122, 139_Alt, 141, 142, 143, 150, 151, 152, 153, 154_W40, 156, 166, 168, 169_Extra, 170, 188, 199, 206, 210, 211, 214, 222, 223, 228, 229, 239, 240, 241, 242
Proportion 1 Transect Training Plots	1 Trans = 19, 4 Trans = 37, Total = 56, Proportion of 1 Trans Plots = 0.339	1 Trans = 18, 4 Trans = 31, Total = 49, Proportion of 1 Trans Plots = 0.367
R^2	0.5223	0.5133
Variables	p50, d8.10m, TotCov	p50, d8.10m, TotCov
Equation	UC = 41.6575 + -7.3908*p50 + -3.3324*d8.10m + 1.7087*TotCov	UC = 48.7549 + -8.3475*p50 + -3.5195*d8.10m + 1.8221*TotCov
Test Plots	92, 230, 108, 151, 188, 168, 212, 42, 139_Alt, 155, 210, 214, 110, 62	303, 90, 121, 184, 155, 27, 230, 42, 80, 243, 31, 198, 107, 108, 247, 167, 106, 212, 79, 105, 207
Proportion 1 Transect Test Plots	1 Trans = 5, 4 Trans = 9, Total = 14, Proportion of 1 Trans Plots = 0.357	1 Trans = 6, 4 Trans = 15, Total = 21, Proportion of 1 Trans Plots = 0.286
Validation R^2	0.3721	0.4739
Validation Equation	y = 51.2434 + 0.3591x	y = 52.2560 + 0.5422x
Min Diff M - P UC	-57.33%	-69.44%
Max Diff M - P UC	85.54%	42.60%
Mean Diff M - P UC	8.38%	-19.44%
Min Abs Diff M - P UC	3.45%	3.28%
Max Abs Diff M - P UC	85.54%	69.44%
Mean Abs Diff M - P UC	37.42%	26.06%

Table 13. Continued

Model	4 Transect Training Plots and 1 Transect Test Plots	Original Model with Plot 121
Training Plots	23, 30, 31, 48, 56, 57, 59, 79, 80, 84, 95, 98, 99, 103, 104, 105, 106, 107, 108, 110, 139_Alt, 141, 142, 143, 154_W40, 155, 156, 166, 167, 168, 169_Extra, 170, 184, 188, 198, 199, 206, 207, 210, 211, 212, 214, 239, 241, 243, 303	13, 14, 16, 26, 27, 62, 79, 80, 84, 89, 90, 92, 95, 98, 99, 103, 104, 105, 106, 107, 108, 121, 122, 141, 142, 143, 150, 151, 152, 153, 168, 210, 211, 212, 214, 222, 223, 228, 229, 230, 239, 240, 241, 242, 243, 247
Proportion 1 Transect Training Plots	1 Trans = 0, 4 Trans = 46, Total = 46, Proportion of 1 Trans Plots = 0.000	1 Trans = 23, 4 Trans = 23, Total = 46, Proportion of 1 Trans Plots = 0.500
R^2	0.4619	0.6547
Variables	p50, d8.10m, TotCov	p50, d8.10m, TotCov
Equation	UC = 40.7542 + -4.7480*p50 + -2.9639*d8.10m + 1.3574*TotCov	UC = 36.5579 + -12.3116*p50 + -9.0397*d8.10m + 2.8275*TotCov
Test Plots	13, 14, 16, 26, 27, 42, 62, 89, 90, 92, 121, 122, 150, 151, 152, 153, 222, 223, 228, 229, 230, 240, 242, 247	23, 30, 31, 42, 48, 56, 57, 59, 110, 139_Alt, 154_W40, 155, 156, 166, 167, 169_Extra, 170, 184, 188, 198, 199, 206, 207, 303
Proportion 1 Transect Test Plots	1 Trans = 24, 4 Trans = 0, Total = 24, Proportion of 1 Trans Plots = 1.000	1 Trans = 1, 4 Trans = 23, Total = 24, Proportion of 1 Trans Plots = 0.042
Validation R^2	0.5828	0.4698
Validation Equation	y = 54.96706 + 0.28842x	y = -73.1089 + 1.5976x
Min Diff M - P UC	-51.66%	-83.55%
Max Diff M - P UC	97.09%	122.67%
Mean Diff M - P UC	8.79%	14.18%
Min Abs Diff M - P UC	0.33%	0.22%
Max Abs Diff M - P UC	97.09%	122.67%
Mean Abs Diff M - P UC	33.47%	46.88%

Table 13. Continued.

was hypothesized that one of the other modeling strategies might produce a better model than the original or the original with plot 121 added back into the training data. This,

again, suggested that more training plots might be necessary to better capture the conditions in Superior National Forest.

The coefficients of determination for the validation data ranged from 0.2547 to 0.5828, with the model built from 4 transect training plots having the highest coefficient of determination. The second highest went to the model built from 70% training data and 30% test, followed by the original model and the original model with plot 121 being tied for the third highest coefficient of determination.

Nearly all of the models had a tendency to under predict the amount of understory cover. Only the stratified random sample model and the 70% training data and 30% test data model over predicted, on average, the understory cover. The 70% training data and 30% test data model had the largest mean difference between the measured and predicted understory cover values for the test plots at -19.44%, while the stratified random sample model was, on average, the closest, with a mean difference of -0.43%. The 80% training data and 20% test data model was also very close on average, with a mean difference between the measured and predicted understory cover of 0.82%.

The mean absolute difference between measured and predicted understory cover values ranged from 25.81% to 47.04%. The original model and the original model with plot 121 were the farthest off the mark, while the 80% training plots and 20% test plots model was the closest, followed by 70% training plots and 30% test plots model.

While the original model and the version with plot 121 added in to the training data preformed pretty well in terms of their coefficients of determination (second and first) and validation coefficients of determination (tied for third), they performed poorly when it came to the looking at the average differences between measured and predicted

understory cover in the validation plots. In terms of mean difference between the measured and predicted amount of understory cover, the models were seventh and sixth, respectively. In terms of the mean absolute difference between the measured and predicted amount of understory cover, the models were eighth and seventh respectively.

The model that had the mean difference between measured and predicted understory cover that was nearest zero was the stratified random sample model. This model also had the third lowest mean absolute difference between measured and predicted understory cover. The stratified random sample model also had the third highest coefficient of determination, but had the third lowest validation coefficient of determination.

The 80% training plots and 20% test plots model had the lowest mean absolute difference between measured and predicted understory cover. This model also had the mean difference between measured and predicted understory cover that was second closest to zero. In terms of the coefficient of determination for the model, it was third worst and second worst in terms of validation coefficient of determination.

The model that was built from plots with 4 transects measured and validated with plots that had one transect measured had the highest validation coefficient of determination, but the second lowest model coefficient of determination. The original model with plot 121 included in the training plots had the highest model coefficient of determination with the proportion of 1 transect plots used to build the model being 0.50. This may have indicated that, in some instances, that the one transect plots were indeed representative of the understory cover variability within the plots.

Summary and Conclusion

Seven additional models were built from the pooled data of all 70 understory plots. The models were built using a variety of strategies including data stratification, random selections of training and testing plots, and using only the plots that had 4 transects measured to construct a model. Comparison of the seven models and the original model showed no clear winner, which indicated that more plot data would likely be necessary for model improvement. Even though no one model stood out as being superior, all of the models were using the same or similar variables. The lidar metrics that were most well correlated with the cover line intercept plot measurements were the height of the 50^{th} percentile of lidar returns, the relative density between 8 and 10 meters, and the total cover.

OVERALL SUMMARY AND CONCLUSIONS

Summary

In an attempt to address management needs in Superior National Forest, an understory cover model was built by combining metrics derived from an airborne lidar point cloud and field data collected using the cover line intercept method. The understory cover model had a reasonably good coefficient of determination and comparison of the model to data from the Monitoring Trends in Burn Severity and LANDFIRE projects increased confidence in the model. Field validation of the model resulted in a correlation coefficient of about 0.69, with a mean difference of 14.3% and a mean absolute difference of 47.0%, when the measured understory cover values were compared to the predicted ones.

The understory cover model was also compared to ancillary data sets. The ancillary data sets included two crown base height models, a forest type model, and land cover data. The crown base height models and the forest type model were produced by Aaron Poznanovic and Michael J. Falkowski of the University of Minnesota Department of Forest Resources. While the land cover data came from the National Land Cover Database. The results of the comparisons indicated a correlation between higher amounts of understory cover and the balsam fir – aspen/paper birch and quaking aspen forest types.

The understory cover model was also compared to, and combined with, data mapping balsam fir basal area. These data were produced by Dr. Peter Wolter of the Iowa State University Department of Natural Resource Ecology and Management. Comparison of the balsam fir basal area data with the understory cover model indicated a

correlation between high amounts of balsam fir and moderate to high amounts of understory cover. The balsam fir data were combined with the understory cover model to produce a balsam fir understory cover model. Validation results were poor for the balsam fir understory cover model, with the model consistently under predicting the balsam fir understory cover.

The original training plots (collected in 2015 and 2016) and the validation plots (collected in 2017) were pooled together. Seven additional understory cover models were built from various combinations of training and test plots generated from the pooled plot data. Comparison of all eight understory cover models showed no clear standout, but that the height of the 50th percentile of lidar returns, the relative density between 8 and 10 meters, and the total cover lidar metrics were consistently being selected by the models.

Conclusions

The conclusions stemming from this research include the following: 1) it appears to be possible to map understory fuels using airborne lidar, however, more plot data may be necessary to improve modelling, since the plot data does not include the entire range of understory cover in Superior National Forest; 2) the lidar metrics of the height of the 50th percentile of lidar returns, the relative density between 8 and 10 meters, and the total cover appear to be correlated with understory fuels; and 3) in this instance, airborne lidar data alone could not be used to identify individual tree species and other data sets need to be combined with the lidar data to assist with the mapping of individual tree species.

While the understory cover model may need some work to improve its predictive power, it performed very well in some instances (Figures 40 and 84). The model

provides managers with the first spatial depiction of understory fuels in Superior National Forest. Armed with this information, Superior National Forest managers have a reasonably good idea of where fire hazards may be high and can plan accordingly. The understory cover model could be used to target areas for mechanical thinning or prescribed fire, intersected with a layer showing the locations of structures within Superior National Forest, or fed into a fuel model.

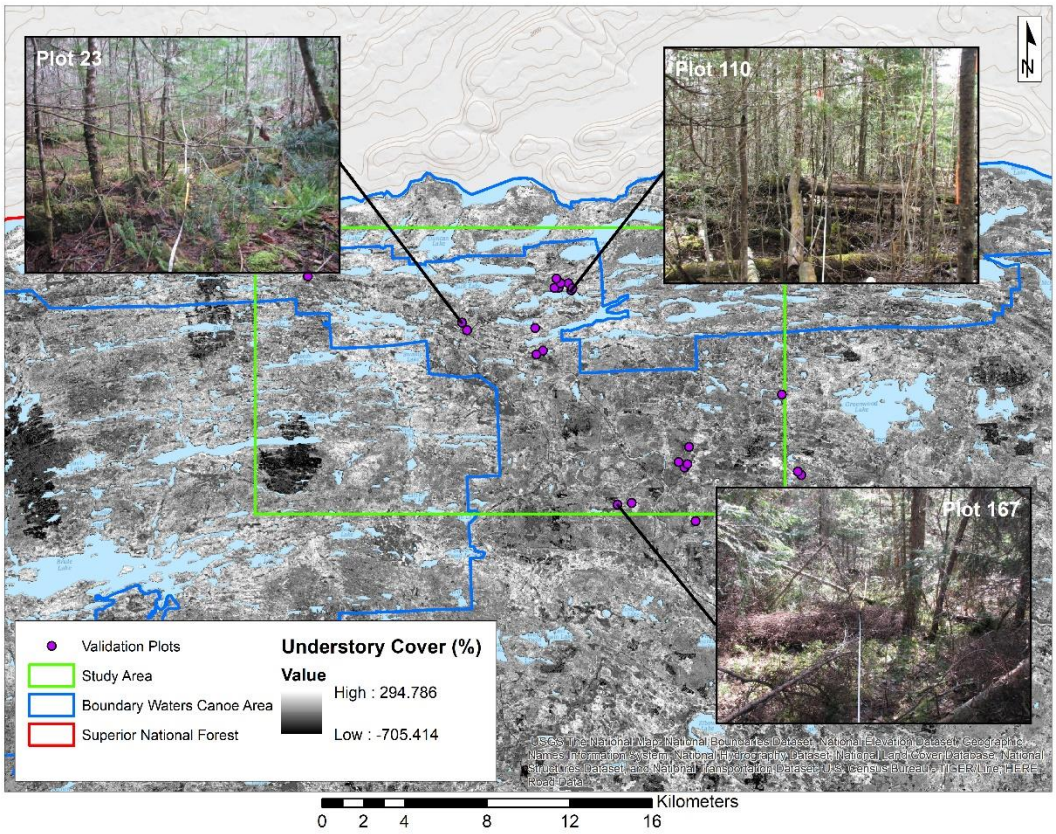

Figure 84. The understory cover model had good predictive power in several validation plots, including plots 23, 110, and 167. Plot 23 had a measured understory cover of 107.37% and a predicted understory cover of 107.87%. Plot 110 had a measured understory cover of 155.51% and a predicted understory cover of 155.13%. Plot 167 had a measured understory cover of 81.12% and a predicted understory cover of 93.44%.

Future Work

In the future, more plot data could be collected within the study area to improve modelling results. After that, the effort could be expanded to other parts of Superior National Forest, until all of Superior National Forest can be mapped with reasonable results. This approach might also work in other forested areas of the nation. Public forests that have had airborne lidar collected might be sought out as areas to further test this approach. Modifications to the approach could be made as needed for each new forest.

REFERENCES

Abbas, Dalia, Dean Current, Mark Ryans, Steven Taff, Howard Horganson & Kenneth N. Brooks. 2011. Harvesting forest biomass for energy - An alternative to conventional fuel treatments: Trials in the Superior National Forest, USA. *Biomass and Bioenergy,* 35**,** 4557-4564.

Andersen, Hans-Erik, Robert J. McGaughey & Stephen E. Reutebuch. 2005. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment,* 94**,** 441-449.

Blais, J. R. 1983a. Predicting Tree Mortality Induced by Spruce Budworm: A Discussion. *The Forestry Chronicle,* 59**,** 294-297.

---. 1983b. Trends in the frequency, extent, and severity of spruce budworm outbreaks in eastern Canada. *Canadian Journal of Forest Research,* 13**,** 539-547.

Boyte, Stephen P., Bruce K. Wylie & Donald J. Major. 2015. Mapping and Monitoring Cheatgrass Dieoff in Rangelands of the Northern Great Basin. *Rangeland Ecology & Management,* 68**,** 18-28.

Brase, Charles H. & Corrinne P. Brase. 2012. *Understandable Statistics: Concepts and Methods, 10th Edition*. Mason, Ohio: Cengage Learning.

Buell, Murray F. & William A. Niering. 1957. Fir-spruce-birch forest in northern Minnesota. *Ecology,* 38**,** 602-610.

Corace III, R. Gregory, Lindsey M. Shartell, Lisa A. Schulte, Wayne L. Brininger, Michelle K. D. McDowell & Daniel M. Kashian. 2012. An Ecoregional Context for Forest Management on National Wildlife Refuges of the Upper Midwest, USA. *Environmental Management,* 49**,** 359-371.

Crawley, Michael J. 2007. *The R Book*. Chichester, West Sussex, England: John Wiley & Sons Ltd.

Demchik, Michael C., Dalia Abbas, Dean Current, Don Arnosti, Myra Theimer & Patty Johnson. 2009. Combining Biomass Harvest and Forest Fuel Reduction in the Superior National Forest, Minnesota. *Journal of Forestry,* 107**,** 235-241.

Esri Inc. 1999-2017. ArcGIS Desktop 10.5.1. 10.5.1.7333.

---. Technical Support: How To: Rescale raster data. Last Modified: 5/5/2016. Esri Inc. Available online: https://support.esri.com/en/technical-article/000008671. Accessed: January 10, 2017

Fleming, Richard A., Jean-Noël Candau & Rob S. McAlpine. 2002. Landscape-scale analysis of interactions between insect defoliation and forest fire in central Canada. *Climatic Change,* 55**,** 251-272.

Fraver, Shawn, Theresa Jain, John B. Bradford, Anthony W. D'Amato, Doug Kastendick, Brian Palik, Doug Shinneman & John Stanovick. 2011. The efficacy of salvage logging in reducing subsequent fire severity in conifer-dominated forests of Minnesota, USA. *Ecological Applications,* 21**,** 1895-1901.

Frelich, Lee E. & Peter B. Reich. 1998. Disturbance Severity and Threshold Responses in the Boreal Forest. *Ecology and Society,* 2.

Friedman, Steven K. & Peter B. Reich. 2005. REGIONAL LEGACIES OF LOGGING: DEPARTURE FROM PRESETTLEMENT FOREST CONDITIONS IN NORTHERN MINNESOTA. *Ecological Applications,* 15**,** 726-744.

Friedman, Steven K., Peter B. Reich & Lee E. Frelich. 2001. Multiple scale composition and spatial distribution patterns of the north-eastern Minnesota presettlement forest. *Journal of Ecology,* 89**,** 538-554.

García, M., D. Riaño, E. Chuvieco, J. Salas & F. M. Danson. 2011. Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. *Remote Sensing of Environment,* 115**,** 1369-1379.

GeoCue Group Inc. 2017. LP360. 2017.1.54.0.

Gilmore, Daniel W., Douglas N. Kastendick, John C. Zasada & Paula J. Anderson. 2003. Alternative Fuel Reduction Treatments in the Gunflint Corridor of the Superior National Forest: Second-year Results and Sampling Recommendations. Research Note NC-381. USDA Forest Service, North Central Research Station, St. Paul, Minnesota, USA.

Goodwin, N. R., N. C. Coops, C. Bater & S. E. Gergel. 2007. Assessment of Sub-Canopy Structure in a Complex Coniferous Forest. *International Archives of Photogrammetry and Remote Sensing,* Volume XXXVI, Part 3/W52**,** 169-172.

Gu, Yingxin & Bruce K. Wylie. 2015. Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations. *Remote Sensing of Environment,* 171**,** 291-298.

---. 2017. Mapping marginal croplands suitable for cellulosic feedstock crops in the Great Plains, United States. *GCB Bioenergy,* 9**,** 836-844.

Gu, Yingxin, Bruce K. Wylie, Stephen P. Boyte, Joshua Picotte, Daniel M. Howard, Kelcy Smith & Kurtis J. Nelson. 2016. An Optimal Sample Data Usage Strategy to Minimize Overfitting and Underfitting Effects in Regression Tree Models Based on Remotely-Sensed Data. *Remote Sensing,* 8.

Heinselman, Miron L. 1973. Fire in the Virgin Forests of the Boundary Waters Canoe Area, Minnesota. *Quaternary Research,* 3**,** 329-382.

---. 1996. *The Boundary Waters Wilderness Ecosystem*. Minneapolis, MN: University of Minnesota Press.

Homer, Collin, Jon Dewitz, Limin Yang, Suming Jin, Patrick Danielson, George Xian, John Coulston, Nathaniel Herold, James Wickham & Kevin Megown. 2015. Completion of the 2011 National Land Cover Database for the Conterminous United States - Representing a Decade of Land Cover Change Information. *Photogrammetric Engineering & Remote Sensing,* 81**,** 345-354.

Isenburg, Martin. 2017. LAStools: award-winning software for efficient LiDAR processing (with LASzip). 171215.

Jones, Robert E. 1968. A Board to Measure Cover Used by Prairie Grouse. *The Journal of Wildlife Management,* 32**,** 28-31.

Kramer, Heather A., Brandon M. Collins, Maggi Kelly & Scott L. Stephens. 2014. Quantifying Ladder Fuels: A New Approach Using LiDAR. *Forests,* 5**,** 1432-1453.

Kramer, Heather A., Brandon M. Collins, Frank K. Lake, Marek K. Jakubowski, Scott L. Stephens & Maggi Kelly. 2016. Estimating Ladder Fuels: A New Approach Combining Field Photography with LiDAR. *Remote Sensing,* 8**,** 766.

LANDFIRE. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST1999. Published: 20120601. Wildland Fire Science, Earth

Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2000. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2001. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2002. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2003. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2004. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2005. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2006. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2007. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2008. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2009. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2010. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2011. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. Disturbance 1999-2014, Fuel Disturbance, Vegetation Disturbance: LANDFIRE.US_DIST2012. Published: 20120601. Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey. Available online: http://landfire.cr.usgs.gov/viewer/. Accessed: November 3, 2016

---. About LANDFIRE (2013, January - last update). Available online: https://www.landfire.gov/about.php. Accessed: February 6, 2018

Lefsky, Michael A., Warren B. Cohen, Geoffrey G. Parker & David J. Harding. 2002. Lidar Remote Sensing for Ecosystem Studies. *BioScience,* 52**,** 19-30.

Lumley, Thomas using Fortran code by Alan Miller. 2009. Leaps: Regression Subset Selection. R package version 2.9. https://CRAN.R-project.org/package=leaps.

Lutes, Duncan C., Robert E. Keane, John F. Caratti, Carl H. Key, Nathan C. Benson, Steve Sutherland & Larry J. Gangi. 2006. FIREMON: Fire effects monitoring and inventory system. General Technical Report RMRS-GTR-164-CD. Fort Collins, CO. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Mendenhall, William M. & Terry L. Sincich. 2016. *Statistics for Engineering and the Sciences Sixth Edition*. Boca Raton, Florida: CRC Press, Taylor & Francis Group.

Menning, Kurt M. & Scott L. Stephens. 2007. Fire Climbing in the Forest: A Semiqualitative, Semiquantitative Approach to Assessing Ladder Fuel Hazards. *Western Journal of Applied Forestry,* 22**,** 88-93.

Minnesota Department of Natural Resources. 2014. LiDAR Elevation, Arrowhead Region, NE Minnesota, 2011. St. Paul, Minnesota: Minnesota Department of Natural Resources. Available: https://gisdata.mn.gov/dataset/elev-lidar-arrowhead2011.

Monitoring Trends in Burn Severity. Data Access: CONUS Thematic Burn Severity Mosaic (2002). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2003). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2004). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2005). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2006). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2007). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2008). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2009). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2010). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2011). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Data Access: CONUS Thematic Burn Severity Mosaic (2012). Published: 20160401. MTBS Project (USDA Forest Service/U.S. Geological Survey). Available online: http://mtbs.gov/dataaccess.html. Accessed: October 24, 2016

---. Project Overview (2017, July - last revised). USDA Forest Service/U.S. Geological Survey. Available online: https://www.mtbs.gov/project-overview. Accessed: February 5, 2018

Mooney, H. A., B. G. Drake, R. J. Luxmoore, W. C. Oechel & L. F. Pitelka. 1991. Predicting Ecosystem Responses to Elevated CO2 ConcentrationsWhat has been learned from laboratory experiments on plant physiology and field observations? *BioScience,* 41**,** 96-104.

Naesset, Erik. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment,* 80**,** 88-99.

Nudds, Thomas D. 1977. Quantifying the Vegetative Structure of Wildlife Cover. *Wildlife Society Bulletin (1973-2006),* 5**,** 113-117.

Poznanovic, Aaron & Michael J. Falkowski. 2015. Final Report: Geospatial Assessment of Forest Fire Fuels in the Boundary Waters Canoe Area. University of Minnesota. Department of Forest Resources.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Renslow, Michael S. 2012. *Manual of Airborne Topographic Lidar*. Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing.

Riaño, David, Erich Meier, Britta Allgöwer, Emilio Chuvieco & Susan L. Ustin. 2003. Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment,* 86**,** 177-186.

Scheller, Robert M., David J. Mladenoff, Thomas R. Crow & Thoedore A. Sickley. 2005. Simulating the Effects of Fire Reintroduction Versus Continued Fire Absence on Forest Composition and Landscape Structure in the Boundary Waters Canoe Area, Northern Minnesota, USA. *Ecosystems,* 8**,** 396-411.

Simonson, W. D., H. D. Allen & D. A. Coomes. 2014. Overstorey and topographic effects on understories: Evidence for linkage from cork oak (Quercus suber) forests in southern Spain. *Forest Ecology and Management,* 328**,** 35-44.

Skowronski, Nicholas, Kenneth Clark, Ross Nelson, John Hom & Matt Patterson. 2007. Remotely sensed measurements of forest structure and fuel loads in the Pinelands of New Jersey. *Remote Sensing of Environment,* 108**,** 123-129.

Stocks, B. J. 1987. Fire Potential in the Spruce Budworm-damaged Forests of Ontario. *The Forestry Chronicle,* 63**,** 8-14.

Sturtevant, Brian R., Brian R. Miranda, Douglas J. Shinneman, Eric J. Gustafson & Peter T. Wolter. 2012. Comparing modern and presettlement forest dynamics of a subboreal wilderness: Does spruce budworm enhance fire risk? *Ecological Applications,* 22**,** 1278-1296.

Swain, Albert M. 1973. A history of fire and vegetation in northeastern Minnesota as recorded in lake sediments. *Quaternary Research,* 3**,** 383-396.

U.S. Geological Survey. 20141010. NLCD 2011 Land Cover (2011 Edition, amended 2014) - National Geospatial Data Asset (NGDA) Land Use Land Cover: None None. U.S. Geological Survey, Sioux Falls, SD. Available online: http://www.mrlc.gov/. Accessed: December 12, 2016

Uchytil, Ronald J. 1991. Abies balsamea. In: Fire Effects Information System, [Online]. U.S. Department of Agriculture. Forest Service. Rocky Mountain Research Station. Fire Sciences Laboratory. Available: http://www.fs.fed.us/database/feis/ [2017, April 20].

United States Department of Agriculture. Superior National Forest - About the Forest. United States Department of Agriculture. Forest Service. https://www.fs.usda.gov/main/superior/about-forest. [2018, January 8]

Wehr, Aloysius & Uwe Lohr. 1999. Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing,* 54**,** 68-82.

Wing, Brian M., Martin W. Ritchie, Kevin Boston, Warren B. Cohen, Alix Gitelman & Michael J. Olsen. 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment,* 124**,** 730-741.

Wolter, Peter T. 2016a. Personal Communication. August 31, 2016. ed. Jeffrey R. Irwin.

---. 2016b. Preliminary Report: Satellite-based forest structure mapping in the Boundary Waters Canoe Area Wilderness for ecosystem management and decision support. Ames, Iowa: Iowa State University. Department of Natural Resource Ecology and Manangement.

---. 2017. Personal Communication. January 12, 2017. ed. Jeffrey R. Irwin.

Wolter, Peter T. & Philip A. Townsend. 2011. Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota. *Remote Sensing of Environment,* 115**,** 671-691.

Wolter, Peter T., Philip A. Townsend & Brian R. Sturtevant. 2009. Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Remote Sensing of Environment,* 113**,** 2019-2036.

Wolter, Peter T., Philip A. Townsend, Brian R. Sturtevant & Clayton C. Kingdon. 2008. Remote sensing of the distribution and abundance of host species for spruce

budworm in Northern Minnesota and Ontario. *Remote Sensing of Environment,* 112**,** 3971-3982.

Wright, Jr Herbert E. & Miron L. Heinselman. 2014. The Ecological Role of Fire in Natural Conifer Forests of Western and Northern North America—Introduction, with an Introduction by Martin E. Alexander. *Fire Ecology,* 10**,** 1-13.

Wylie, B. K., L. Zhang, N. Bliss, L. Ji, L. L. Tieszen & W. M. Jolly. 2008. Integrating modelling and remote sensing to identify ecosystem performance anomalies in the boreal forest, Yukon River Basin, Alaska. *International Journal of Digital Earth,* 1**,** 196-220.